

C:Accession: B70888
 R:Colo, S.T.; Broesh, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, L.; Gordon, S.; Connor, P.; Davies, P.; Devlin, F.; Felwell, T.; Genot, S.; Hamlin, N.; Holtz, S.; Rastam, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellern, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: 49836, P.; Sultana, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; M01098;95987; PMID:9644230
 A:Accession: F70888
 A>Status: preliminary (unpublished sequence not shown, translated not shown)
 A:Molecule type: DNA
 A:Residues: 1-302 <CDS>
 A:Cross-references: GR: A1022076; GR: A123456; NID:94256025; EMBL: S6A17470.1; F10-0125569
 A:Experimental source: strain H379
 C:Genetics:
 A:Gene: Rv0800c
 A:Map position:

Query Match 100.0% Score 25; DB 2; Length 302;
 Best Local Similarity 100.0% Pred. No. 1 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 1111
 Db 210 LESYT 214

RESULT 13
 64388
 hypothetical protein M0711 - *Methanococcus jannaschii*

C:Species: *Methanococcus jannaschii*
 C:Date: 14-Sep-1996 #sequence_revision 11-Sep-1996 #text_change 21-Jul-2000
 C:Accession: 64388
 R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, J.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overback, R.; Kinknes, E.P.; Wainstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Slick, F.W.; Hickey, M.D.; Gelfand, M.D.; Sabetis, K.M.; Huft, M.A.; Science 273, 1068-1073, 1996
 A:Authors: Kline, R.P.; Poroskany, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
 A:Reference number: A9300; M01096;93799; PMID:8688087
 A:Accession: 64388
 A>Status: preliminary (unpublished sequence not shown, translation not shown)
 A:Molecule type: DNA
 A:Residues: 1-322 <CDS>
 A:Cross-references: GR: 066759; GR: 077117; NID:9292712; EMBL: AAR0014.1; F10-0159129
 C:Genetics:
 A:Map position: F064397-64495

Query Match 100.0% Score 25; DB 2; Length 322;
 Best Local Similarity 100.0% Pred. No. 1 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 1111
 Db 106 LESYT 110

RESULT 14
 A96596
 hypothetical protein T1812 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A96596
 P:Thole, A.; Ecker, J.P.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chen, C.W.; Chung, M.K.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, R.; Huizar, L.; Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, F.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, F.; Rowley, D.; Sakano, H.; A:Authors: Salter, S.; Schmitz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A96141; M01021016719; PMID:1110712

A:Accession: A96596
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-354 <STO>
 A:Cross-references: GR: A905173; NID:11094798; EMBL: AAC29720.1; GSPDB: GNO0141
 C:Genetics:
 A:Gene: T1812
 A:Map position: 1

Query Match 100.0% Score 25; DB 2; Length 354;
 Best Local Similarity 100.0% Pred. No. 1 4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 1111
 Db 315 LESYT 319

RESULT 15
 T08277
 carotenoid biosynthesis protein homolog H0660 - *Halobacterium* sp. (strain NRC-1) plas

N:Alternate names: hypothetical protein H1696
 C:Species: *Halobacterium* sp.
 A:Variety: strain NRC-1
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 31-Jan-2000
 C:Accession: T08277; T08366
 R:Ng, W.V.; Clift, S.A.; Smith, I.M.; Rungtner, R.E.; Paskin, D.; Faust, J.; Hall, B.; Genome Res. 8, 1131-1141, 1998
 A>Title: Snapshot of a large dynamic replicon in a halophilic *Archaeon*: megaplasmid
 A:Reference number: Z16408; M01096;9795; PMID:9847077
 A:Accession: T08277
 A>Status: translated from GR/EMBL/AF183
 A:Molecule type: DNA
 A:Residues: 1-379 <NCM>
 A:Cross-references: EMBL: AF01649; NID:9282279; F10-0282338; HALOSP:H0660
 A:Experimental source: strain NRC-1
 A:Genetics: CDP1
 A:Accession: T08366
 A>Status: translated from GR/EMBL/CDR1
 A:Molecule type: DNA
 A:Residues: 1-379 <DAS>
 A:Cross-references: EMBL: AF01649; NID:9282279; F10-0282327; HALOSP:H1696
 A:Experimental source: strain NRC-1
 A:Genetics: CDP2
 C:Genetics: <CDP1>
 A:Gene: HALOSP:H0660
 A:Gene: HALOSP:H0660
 C:Genetics: <CDP2>
 A:Gene: HALOSP:H1696
 A:Gene: plasmid pNRC100
 C:Superfamily: carotenoid biosynthesis protein homolog

Query Match 100.0% Score 25; DB 2; Length 379;
 Best Local Similarity 100.0% Pred. No. 1 5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 1111
 Db 131 LESYT 135

Search completed: April 8, 2004, 17:05:09
 Job time: 26.0833 secs

Accession: S44281
 C:Spulman, F.; Salamin, F.
 Submitted to the EMBL Data Library, February 1994
 A:Reference number: S44281
 A:Accession: S44281
 A:Molecule type: mRNA
 A:Residues: 1-227 <SPS>
 A:Cross-references: EMBL:X7575; NID:g442918; PID:g1345528
 C:Genetics:
 A:Gene: EST

Query Match
 Best Local Similarity 100.0% Score 25; DB 2; Length 227;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
 DB 204 LEAST 208

RESULT 8

probable 1-acylglycerol-3-phosphate O-acyltransferase (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Apr-1994 #sequence_revision 20-Aug-1994 #text_change 20-Aug-1999
 C:Accession: G72484
 C:Keywords: alkyltransferase; Y1; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatayama, Y.; Jinno, K.; Takai, H.; Takamata, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; R
 DNA Res. 6: 44-50, 1999
 A:Title: The complete genome sequence of an aerobic hyperthermophilic Gram-negative, Aeropyr
 A:Reference number: A74501; M01D:99310349; PMID:10382966
 A:Accession: G72484
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-235 <KAM>
 A:Cross-references: DDBJ:AF000064; NID:g5105945; PIRN:BAA81535.1; PID:d1045321; PTD:g510
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2519

Query Match
 Best Local Similarity 100.0% Score 25; DB 2; Length 235;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
 DB 207 LEAST 211

RESULT 9

probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori
 C:Species: Helicobacter pylori
 A:Variant: strain 199
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 06-Oct-2000
 C:Accession: B71827
 C:Keywords: R.A.; Ling, L.S.L.; Mofit, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 ; Ives, C.; Gibson, F.; Mettrey, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voytas, G.F.;
 Nature 397: 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; M01D:99120557; PMID:9936932
 A:Accession: B71827
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-237 <ARN>
 A:Cross-references: GR:AE001550; GR:AE001439; NID:g4155872; PIDN:AD06952.1; PTD:g415587
 A:Experimental source: strain 199
 C:Genetics:
 A:Gene: p152
 C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase
 C:Keywords: acyltransferase; coenzyme A

Query Match 100.0% Score 25; DB 2; Length 237;

Best Local Similarity 100.0% Pred No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
 DB 204 LEAST 208

RESULT 10

probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter py
 C:Species: Helicobacter pylori
 C:Date: 09-Apr-1997 #sequence_revision 09-Apr-1997 #text_change 06-Oct-2000
 C:Accession: D64688
 C:Keywords: R.Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
 F.; White, P.; Adams, R.; Holt, C.; Peterson, J.; Smith, R.; Venter, A.; M. B.
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Wadman, J.M.; Fujita, O.; Bowman, C.; Wathey,
 Nature 388: 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karik, P.; Smith, H.O.; Fraser,
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; M01D:97594467; PMID:9252185
 A:Accession: D64688
 A:Status: Preliminary, nucleic acid sequence not shown, translation not shown
 A:Molecule type: DNA
 A:Residues: 1-240 <TOM>
 A:Cross-references: GR:AF000636; GR:AF000511; NID:g2414517; PIRN:CA08393.1; PTD:g241
 C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase
 C:Keywords: acyltransferase; coenzyme A

Query Match
 Best Local Similarity 100.0% Score 25; DB 2; Length 240;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
 DB 204 LEAST 208

RESULT 11

probable integral membrane protein (imported) Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 25-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: G86920
 C:Keywords: R.Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
 R.; Davies, R.M.; Devlin, K.; Doherty, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Hollo
 eam, M.A.; Rutherford, K.M.
 Nature 409: 1007-1011, 2001
 A:Authors: Rutherford, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skellern, J.; Squares, P.;
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; M01D:01128732; PMID:11234092
 A:Accession: G86920
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <STO>
 A:Cross-references: GR:A450380; NID:g13092477; PIRN:CA229603.1; GSTRB:GN0147
 C:Genetics:
 A:Gene: MD0095

Query Match
 Best Local Similarity 100.0% Score 25; DB 2; Length 402;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEAST 5
 DB 210 LEAST 214

RESULT 12

hypothetical protein Rv3806c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1996 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 48 LESYT 52

RESULT 4

E90199 conserved hypothetical protein (imported) - *Sulfolobus solfataricus*

C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: E90199

C:Author: Singh, R.K.; Conlath, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan, J.; Jettles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.

C:Submitted to Genbank, April 2001

C:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: E90199

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-133 <RUP>

A:Cross-references: GB:AE00641; MID:q12813694; PIRN:AAK40652; GSPDR:GN00155

C:Genetics:

C:Superfamily: conserved hypothetical protein MJ1552

Query Match 100.0%; Score 25; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4

T02619 hypothetical protein At2g26040 (imported) - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001

C:Accession: T02619; G84055

C:Author: R. Kaul, S. Li, X. Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, August 1998

C:Description: *Arabidopsis thaliana* chromosome II BAC T19L8 genomic sequence.

A:Reference number: Z14681

A:Accession: T02619

A:Status: translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-190 <RUP>

A:Cross-references: EMBL:AC004747; NID:q4133696; PIR:q413709

A:Experimental source: cultured Columbia

A:Author: R. Kaul, S. Li, Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;

M. Koo, H. Motil, R. S. Conrath, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; M010:20083487; PMID:10617197

A:Accession: G84655

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-190 <STO>

A:Cross-references: GR:AE002093; NID:q413709; PIR:AC013321; GSPDR:GN00139

C:Genetics:

A:Gene: T19L18.15; At2g26040

A:Map position: 2

C:Superfamily: *Arabidopsis thaliana* hypothetical protein T1358.30

Query Match 100.0%; Score 25; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 74;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 146 LESYT 150

RESULT 5

E90187 HAM1 protein (imported) - *Sulfolobus solfataricus*

C:Species: *Sulfolobus solfataricus*

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: E90187

C:Author: Singh, R.K.; Conlath, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan,

J.; Jettles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,

submitted to Genbank, April 2001

C:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: E90187

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-192 <RUP>

A:Cross-references: GB:AE00641; MID:q12813694; PIRN:AAK40756; GSPDR:GN00155

C:Genetics:

C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0226

Query Match 100.0%; Score 25; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6

AH2547 hypothetical protein alr7667 (imported) - *Nostoc* sp. (strain PCC 7120) plasmid pCC7

A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C:Species: *Nostoc* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AH2547

C:Author: R. Kaneko, T.; Nakamura, Y.; Maki, C.P.; Kurita, T.; Sasagawa, S.; Watanabe, A.; Iri,

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Taka,

DNA Res. 8, 205-213, 2001

A:Title: Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium

A:Reference number: AB1807; M010:2159285; PMID:11759840

A:Accession: AH2547

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <RUP>

A:Cross-references: GR:AP003602; PIRN:BAW77410.1; PIR:q17144752; GSPDR:GN00181

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr7667

A:Gene: alr7667

A:Gene: plasmid

Query Match 100.0%; Score 25; DB 2; Length 216;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
|||||
DB 147 LESYT 151

RESULT 7

S44281 E543 protein - barley

C:Species: *Hordeum vulgare* (barley)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Mar-1999

GenCore version 5.1.4-PS-4578
Copyright (c) 1993 2003 CompuGen Ltd.

3M protein - protein search using sw model

Run on: April 8, 2003, 11:09:15 : Search time 22.943 seconds
(without alignments)
21766 Million cell updates/sec

Title: US-09-647-749A-1

Perfect score: 25

Sequence: 1 LEVST 5

Scoring table: RUSUM2

Gapop 10.0, Gapext 0.5

Searched: 281224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%
Maximum first 45 summaries

Database:

1: PIR_731*
2: PIR1*
3: PIR2*
4: PIR3*
5: PIR4*

Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 100.0 | 99 | 2 | JC5295 | monocyte chemotact |
| 2 | 100.0 | 99 | 2 | JC2417 | monocyte chemotact |
| 3 | 100.0 | 133 | 2 | E90199 | conserved hypotet |
| 4 | 100.0 | 190 | 2 | T02619 | hypothetical prote |
| 5 | 100.0 | 192 | 2 | E90187 | HAM1 protein (limp |
| 6 | 100.0 | 216 | 2 | AH2547 | hypothetical prote |
| 7 | 100.0 | 227 | 2 | S44281 | ES43 protein - bar |
| 8 | 100.0 | 245 | 2 | G72484 | hypothetical prote |
| 9 | 100.0 | 247 | 2 | B71827 | probable 1-acylaly |
| 10 | 100.0 | 240 | 2 | D64688 | probable 1-acylaly |
| 11 | 100.0 | 302 | 2 | G86920 | probable integral |
| 12 | 100.0 | 302 | 2 | B70888 | hypothetical prote |
| 13 | 100.0 | 352 | 2 | G64388 | hypothetical prote |
| 14 | 100.0 | 354 | 2 | A96596 | hypothetical prote |
| 15 | 100.0 | 379 | 2 | T08277 | catenoid biosynt |
| 16 | 100.0 | 396 | 2 | F87647 | phosphoserine amin |
| 17 | 100.0 | 454 | 2 | T39297 | probable atpase - |
| 18 | 100.0 | 458 | 2 | E97349 | membrane associate |
| 19 | 100.0 | 442 | 2 | G01804 | interleukin 3-reou |
| 20 | 100.0 | 470 | 2 | D41977 | retinoid receptor |
| 21 | 100.0 | 532 | 2 | D71267 | probable UDP-N-ace |
| 22 | 100.0 | 548 | 2 | T16082 | hypothetical prote |
| 23 | 100.0 | 597 | 2 | D82761 | lipase/esterase MF |
| 24 | 100.0 | 634 | 2 | A25646 | dnak-type molecula |
| 25 | 100.0 | 633 | 2 | A55719 | dnak-type molecula |
| 26 | 100.0 | 739 | 2 | B86606 | primosomal protein |
| 27 | 100.0 | 749 | 2 | G72018 | primosomal protein |
| 28 | 100.0 | 753 | 2 | R71472 | probable primosoma |
| 29 | 100.0 | 753 | 2 | D81734 | primosomal protein |

| | | | | | |
|----|----|-------|------|--------|--------------------|
| 30 | 25 | 100.0 | 756 | S47656 | LMC II protein - |
| 31 | 25 | 100.0 | 872 | JC7380 | RNA-directed RNA p |
| 32 | 25 | 100.0 | 922 | AC1827 | matricoplytreha |
| 33 | 25 | 100.0 | 1157 | B69163 | DNA helicase relat |
| 34 | 25 | 100.0 | 1252 | S77037 | hypothetical prote |
| 35 | 25 | 100.0 | 1679 | T30271 | surface protein - |
| 36 | 25 | 100.0 | 2004 | AC0314 | probable membrane |
| 37 | 25 | 100.0 | 2397 | A55335 | protein precursor |
| 38 | 25 | 100.0 | 2409 | A66979 | versican precursor |
| 39 | 25 | 100.0 | 2543 | F69679 | polyketide synthas |
| 40 | 25 | 100.0 | 2871 | A55624 | thirillin-1 precur |
| 41 | 25 | 100.0 | 5825 | 112117 | polypeptide - fava |
| 42 | 23 | 92.0 | 51 | B70141 | hypothetical prote |
| 43 | 23 | 92.0 | 112 | AC0661 | hypothetical prote |
| 44 | 23 | 92.0 | 119 | T08158 | hypothetical lumen |
| 45 | 23 | 92.0 | 117 | T08271 | probable thiorodox |

ALIGNMENTS

RESULT 1
JC5295
monocyte chemotactic protein-2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 02-May-1997 #sequence: J01510.19-741-1997 #rev: _chango 20-Jun-2000
C:Accession: JC5295
R:Van Gollie, E.; Froyen, G.; Namiyama, H.; Miura, K.; Piten, P.; Van Aelst, L.; W
Blochem. Biophys. Res. Commun. 211, 720-730, 1997
A>Title: Human monocyte chemotactic protein-2: cDNA cloning and regulated expression
A:Reference number: JC5295; M01199224420; PMID:9070881
A:Accession: JC5295
A:Molecule type: mRNA
A:Residues: 1-99 <VAN>
A:Cross-references: GB:Y10802; NID:01924937; PIDN:CAAT7160.1; PID:01924938
A:Experimental source: bone marrow
A:Comment: This protein belongs to the beta-chemokine family which is one of the ma
tis and in tumor biology, and contribute to the trafficking and recruitment of the
C:Genetics:
A:Gene: mcp-2
C:Superfamily: macrophage inflammatory protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-99/Product: monocyte chemotactic protein-2 #status predicted <MAT>
Query Match 100.0% Score 25; DB 2; Length 99;
Best local similarity 100.0% Pred No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LEVST 5
| | | | |
Db 48 LEVST 52
| | | | |
RESULT 2
JC2417
monocyte chemotactant protein-2 precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 24-Feb-1995 #sequence: J01510.24-Feb-1995 #text_change: 16-Jul-1999
C:Accession: JC2417
R:Hosang, K.; Knöke, I.; Klading, J.; Wenne, F.; Wille, W.; Schmitt, K.H.
Biochem. Biophys. Res. Commun. 206, 148-154, 1994
A>Title: Porcine leucal cells express monocyte chemotactant protein-2 (MCP-2): At
A:Reference number: JC2417; M0119504716; PMID:799015
A:Accession: JC2417
A:Molecule type: mRNA
A:Residues: 1-99 <HOS>
A:Cross-references: GB:249480; NID:01924938; PIDN:CAAT8471.1; PID:01924939
A:Experimental source: corpus luteum
A:Superfamily: macrophage inflammatory protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-99/Product: monocyte chemotactant protein-2 #status predicted <MAT>
Query Match 100.0% Score 25; DB 2; Length 99;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using SW model

```
Run on: April 8, 2003, 11:55:39 ; Search time 43.3333 Seconds
        (without alignments)
        15.375 Million cell updates/sec
```

Title: US-09-647-749A-1

Perfect score

Sequence

Normative

jaapop 10.0 , jaapext 0.5

Search: 108470 seqs, 133250620 residues

```

Minimum Fr seq length: 0
Maximum Fr seq length: 200000000

```

| | |
|------------------|--------------------|
| Post-processing: | Minimum Match 08 |
| | Maximum Match 1000 |

Maximum match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:

1: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1380.DAT
2: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1391.DAT
3: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1392.DAT
4: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1393.DAT
5: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1394.DAT
6: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1395.DAT
7: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1396.DAT
8: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1397.DAT
9: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1398.DAT
10: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1399.DAT
11: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1400.DAT
12: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1401.DAT
13: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1399.DAT
14: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1399.DAT
15: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1394.DAT
16: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1394.DAT
17: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1394.DAT
18: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1397.DAT
19: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1398.DAT
20: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1399.DAT
21: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1399.DAT
22: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1394.DAT
23: */SLD3/gcd3a/gene5eq/gene5eqp.embl*, AA1394.DAT

Print. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|---------------------|
| 1 | 25 | 100.0 | 5 | 20 | AAV49695 | HIV-1 gp120 inducer |
| 2 | 25 | 100.0 | 50 | 22 | AAU56783 | Propionibacterium |
| 3 | 25 | 100.0 | 58 | 22 | AAW4554 | Human brain cytosol |
| 4 | 25 | 100.0 | 71 | 20 | AAV07234 | Truncated monocytic |
| 5 | 25 | 100.0 | 71 | 20 | AAV07228 | Truncated monocytes |
| 6 | 25 | 100.0 | 72 | 16 | AAV70804 | Chemottractant MO |
| 7 | 25 | 100.0 | 76 | 21 | AAV65931 | Amino acid sequenc |
| 8 | 25 | 100.0 | 76 | 23 | AAU21099 | Protein of WT-MCP- |
| 9 | 25 | 100.0 | 76 | 23 | AAU21105 | MTF-2 chromatin f1 |
| 10 | 25 | 100.0 | 76 | 23 | AAV78390 | Human chromatin HR |

| | | | | | |
|----|----|-------|--------|-----------|----------------------|
| 1 | 25 | 100.0 | 76.4 | AAB56453 | Human monocytic cell |
| 12 | 25 | 100.0 | 77.2 | AAB515785 | Human monocytic MC |
| 13 | 25 | 100.0 | 77.4 | AAB460903 | Human MCF2 triamphen |
| 14 | 25 | 100.0 | 77.2 | AAC749305 | Mouse/human hybrid |
| 15 | 25 | 100.0 | 77.2 | AAG683358 | Chimeric chemokine |
| 16 | 25 | 100.0 | 74.4 | AA174844 | Mouse/human hybrid |
| 17 | 25 | 100.0 | 94.4 | AAAG59356 | Chimeric chemokine |
| 18 | 25 | 100.0 | 95.2 | ABB119228 | Human chemokine |
| 19 | 25 | 100.0 | 95.26 | AAAY05300 | C-3 chemokine, MCP-1 |
| 20 | 25 | 100.0 | 99.2 | AAAY02294 | Wild type monocyte |
| 21 | 25 | 100.0 | 99.26 | AAAY07237 | Wild type monocyte |
| 22 | 25 | 100.0 | 109.18 | AAAW25655 | Human beta-chemokil |
| 23 | 25 | 100.0 | 109.9 | AAWA42072 | Human M6 proprotein |
| 24 | 25 | 100.0 | 109.24 | AAAG36167 | Reference sequence |
| 25 | 25 | 100.0 | 109.23 | AAAG36168 | Polymorphic variant |
| 26 | 25 | 100.0 | 109.23 | AAAG36169 | Polymorphic variant |
| 27 | 25 | 100.0 | 151.22 | AAAU17310 | Novel signal trans |
| 28 | 25 | 100.0 | 209.22 | AAU025707 | G protein-coupled |
| 29 | 25 | 100.0 | 229.2 | AAE759669 | Lactium peptide type |
| 30 | 25 | 100.0 | 240.16 | AAAS55540 | H. pylori ORF 1240 |
| 31 | 25 | 100.0 | 240.19 | AAAG98437 | H. pylori GHO 9034 |
| 32 | 25 | 100.0 | 242.18 | AAAS55441 | H. pylori ORF 06ap1 |
| 33 | 25 | 100.0 | 257.41 | AAAG07701 | Aradidopsis thaliana |
| 34 | 25 | 100.0 | 268.22 | AAAG38392 | Novel human diamon |
| 35 | 25 | 100.0 | 282.20 | AAAY48330 | Human prostate can |
| 36 | 25 | 100.0 | 301.20 | AAAW46031 | BS1 restriction en |
| 37 | 25 | 100.0 | 332.22 | AAAG59521 | Orphan actin from |
| 38 | 25 | 100.0 | 332.22 | AAAB79036 | Corynebacterium gl |
| 39 | 25 | 100.0 | 348.21 | AAAG07700 | Arabidopsis thaliana |
| 40 | 25 | 100.0 | 354.21 | AAAG07699 | Arabidopsis thaliana |
| 41 | 25 | 100.0 | 374.22 | AAAW49973 | Human polypeptide |
| 42 | 25 | 100.0 | 385.22 | AAAG06064 | Novel human diamon |
| 43 | 25 | 100.0 | 445.22 | AAAG36187 | Human polypeptide |
| 44 | 25 | 100.0 | 467.22 | AAAB73685 | Human oxidoreducta |
| 45 | 25 | 100.0 | 477.21 | AAAT21522 | Human dehydrogenase |

AI COMMENTS

RESULT
AAY49695

to standard point 1.5 AA

AC MAY 4 1965

18-744-2000 (first entry.)

Fig. 1. α 20 induced neur and cell death inhibiting peptide #1

PK Nucleocapsid: detection: HIV-1: envelope protein: gp120

KW neuronal cell death: encephalopathy: neuropathy: memory loss:

KW neurotoxicity; inflammatory neurological disease; multiple sclerosis

XX

XX

XX

XX

1. ☒ 2. ☒ 3. ☒ 4. ☒ 5. ☒ 6. ☒ 7. ☒ 8. ☒ 9. ☒ 10. ☒ 11. ☒ 12. ☒ 13. ☒ 14. ☒ 15. ☒ 16. ☒ 17. ☒ 18. ☒ 19. ☒ 20. ☒ 21. ☒ 22. ☒ 23. ☒ 24. ☒ 25. ☒ 26. ☒ 27. ☒ 28. ☒ 29. ☒ 30. ☒ 31. ☒ 32. ☒ 33. ☒ 34. ☒ 35. ☒ 36. ☒ 37. ☒ 38. ☒ 39. ☒ 40. ☒ 41. ☒ 42. ☒ 43. ☒ 44. ☒ 45. ☒ 46. ☒ 47. ☒ 48. ☒ 49. ☒ 50. ☒ 51. ☒ 52. ☒ 53. ☒ 54. ☒ 55. ☒ 56. ☒ 57. ☒ 58. ☒ 59. ☒ 60. ☒ 61. ☒ 62. ☒ 63. ☒ 64. ☒ 65. ☒ 66. ☒ 67. ☒ 68. ☒ 69. ☒ 70. ☒ 71. ☒ 72. ☒ 73. ☒ 74. ☒ 75. ☒ 76. ☒ 77. ☒ 78. ☒ 79. ☒ 80. ☒ 81. ☒ 82. ☒ 83. ☒ 84. ☒ 85. ☒ 86. ☒ 87. ☒ 88. ☒ 89. ☒ 90. ☒ 91. ☒ 92. ☒ 93. ☒ 94. ☒ 95. ☒ 96. ☒ 97. ☒ 98. ☒ 99. ☒ 100. ☒

[illegible]

XX

[illegible]

XX

5

PT New peptides useful for inhibiting human immunodeficiency virus type 1
 PT (HIV-1) gp120 induced neuronal cell death
 XX
 XX
 PS Claim 1: Page 11; 16pp; English.
 XX
 CC AAY49695 and AAY49696 represent peptides which inhibit HIV-1 gp120
 CC induced neuronal cell death. Pharmaceutical compositions containing the
 CC peptides are useful for treating symptoms caused by neuronal cell loss.
 CC Such conditions especially associated with HIV infection include:
 CC encephalopathies, neuropathies, memory loss, dementia, depression,
 CC psychosis and opportunistic infections. The peptides act as antagonists
 CC of HIV-mediated neurotoxicity and subsequent neuronal degeneration.
 CC This enables therapeutic treatment of HIV infection and other
 CC inflammatory neurological diseases, including multiple sclerosis.
 CC tropical spastic paraparesis and Alzheimer's disease.
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 25; DB 23; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESTY 5
 DB 1 LESTY 5
 RESULT 2
 ID AAY56783
 ID AAY56783 standard; Protein; 50 AA.
 XX
 AC AAY56783;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #17679.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN W0200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 20-APR-2001; 2001WC-US12865.
 XX
 21-APR-2001; 2000US-19947P.
 XX
 07-JUN-2000; 2000US-00841P.
 XX
 07-JUL-2000; 2000US-215747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Alaisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 XX
 DR N-PSDB; AAS59578.
 XX
 FT Propionibacterium acnes polypeptides and nucleic acids useful for
 FT vaccinating against and diagnosing infections, especially useful for
 FT treating acne vulgaris -
 PT
 XX
 PS Example 1: SEQ ID NO 17978; 1059pp; English.
 XX
 CC Sequences AAY39105-AAY68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC the disease. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis

CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence. For example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 50 AA;
 Query Match 100.0%; Score 25; DB 22; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESTY 5
 DB 14 LESTY 18
 RESULT 3
 ID AAM64554
 ID AAM64554 standard; Protein; 68 AA.
 XX
 AC AAM64554;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SE; ID NO: 36659.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN W0200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 30-JAN-2001; 2001WO-US06667.
 XX
 04-FEB-2000; 2000US-018612.
 XX
 25-MAY-2000; 2000US-020156.
 XX
 08-JUN-2000; 2000US-060816.
 XX
 03-AUG-2000; 2000US-063445.
 XX
 21-SEP-2000; 2000US-0234687.
 XX
 27-SEP-2000; 2000US-0236687.
 XX
 04-OCT-2000; 2000US-0024763.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 FT Single exon nucleic acid probes for analyzing gene expression in human
 FT brains -
 PT
 XX
 PS Example 4: SEQ ID NO: 36659; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX
 SQ Sequence 68 AA:

Query Match 100.0% Score 25; DB 22; Length 68;
 Best Local Similarity 100.0% Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 11111
 DB 25 LESYT 29

RESULT 4
 AAY07234
 ID AAY07234 standard; protein: 71 AA.

XX AAY07234;
 AC
 XX 06-JUL-1999 (first entry)

DE Truncated monocyte chemotactic protein 2 (6-76)

XX
 KM Wild type: C-C chemokine; monocyte chemotactic protein 2; MCP2; HIV;
 KM regulated on activation normal T-cell expressed and secreted; RANTES;
 KM truncation; antagonist; medicaments; diagnosis; inflammation; infection;
 KM tumour; angiogenesis; hematopoiesis; autoimmune disease; atherosclerosis;
 KM pulmonary disease; skin disorder.

XX Homo sapiens.
 OS Synthetic.

XX EP906954-A1.

XX 07-APR-1999.

XX 29-SEP-1997; 97EP-0116863.

XX 29-SEP-1997; 97EP-0116863.

XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV

XX Proost P, Struyt S, Van Damme J.

XX WPI: 1999-207109/18.

XX New amino-terminally truncated C-C chemokines have antagonistic
 PT activity for treatment of immune, inflammatory and infectious
 PT diseases

XX Claim 4; Fig 1; 29pp; English.

XX This sequence represents a truncated C-C chemokine monocyte chemotactic
 CC protein 2 (MCP2) containing amino acids 6-76 of the mature protein.
 CC The invention relates the generation of amino-terminal truncated C-C
 CC chemokines, having chemokine antagonistic activity. The new chemokines
 CC are useful as medicaments, for diagnosis and/or treatment of diseases
 CC which require antagonistic activity of a chemokine e.g. inflammatory
 CC diseases, HIV infection, tumours, and angiogenesis and hematopoiesis
 CC related diseases, including auto-immune diseases, atherosclerosis,
 CC pulmonary diseases and skin disorders.

XX Sequence 71 AA:

Query Match 100.0% Score 25; DB 20; Length 71;
 Best Local Similarity 100.0% Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 11111
 DB 20 LESYT 24

RESULT 5
 AAY07238
 ID AAY07238 standard; protein: 71 AA.

XX AAY07238;

XX 06-JUL-1999 (first entry)

DE Truncated monocyte chemotactic protein 2 (6-76).

XX Wild type: C-C chemokine; monocyte chemotactic protein 2; MCP2; HIV;
 KM regulated on activation normal T-cell expressed and secreted; RANTES;
 KM truncation; antagonist; medicaments; diagnosis; inflammation; infection;
 KM tumour; angiogenesis; hematopoiesis; autoimmune disease; atherosclerosis;
 KM pulmonary disease; skin disorder.

XX Homo sapiens.
 OS Synthetic.

XX EP905241-A1.

XX 31-MAR-1999.

XX 10-MAR-1998; 98EP-0104215.

XX 19-DEC-1997; 97EP-0122471.

XX 29-SEP-1997; 97EP-0116863.

XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

XX Proost P, Struyt S, Van Damme J.

XX WPI: 1999-206774/18.

XX New amino-terminally truncated C-C chemokines have antagonistic
 PT activity for treatment of immune, inflammatory and infectious
 PT diseases

XX Claim 4; Fig 1; 29pp; English.

XX This sequence represents a truncated C-C chemokine monocyte chemotactic
 CC protein 2 (MCP2) containing amino acids 6-76 of the mature protein.
 CC The invention relates the generation of amino-terminal truncated C-C
 CC chemokines, having chemokine antagonistic activity. The new chemokines
 CC are useful as medicaments, for diagnosis and/or treatment of diseases
 CC which require antagonistic activity of a chemokine e.g. inflammatory
 CC diseases, HIV infection, tumours, and angiogenesis and hematopoiesis-
 CC related diseases, including auto-immune diseases, atherosclerosis,
 CC pulmonary diseases and skin disorders.

XX Sequence 71 AA:

Query Match 100.0% Score 25; DB 20; Length 71;
 Best Local Similarity 100.0% Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
 11111
 DB 20 LESYT 24

RESULT 6
 AAR70804
 ID AAR70804 standard; protein: 72 AA.

XX AAR70804;

XX 29-AUG-1995 (first entry)

DE Chemottractant MCP-2.

XX Chemottractant; MCP-2; heparanase; heparin; heparan sulfate;

KW arthritis; restenosis; cancer; wound healing.
 XX Homo sapiens.
 US W0504158-A.
 PN 09-FEB-1995.
 PD 26-JUL-1994: 94WO-US08207.
 XX 29-JUL-1993: 94US-0099866.
 PR 13-OCT-1993: 93US-0136117.
 XX (UPJO) UPJOHN CO.
 PA
 XX
 PI Hooywerf AJ, Ledbetter SR;
 XX WPI: 1995-092239/11.
 DR
 XX Screening for cFds. with anti-heparanase activity - by detecting
 inhibition of heparin or heparan sulphate degradation,
 potentially useful for treating arthritis, restenosis, cancer.
 PT Claim 13: Page 53: 60pp; English.
 PS
 XX Purified heparanases, prepared under reducing conditions and
 activated with transglutaminase, are given in AAR70786-804. Most
 are prepared by reverse transcription of mRNA from activated human
 CC leukocytes, then cloning of the cDNA into pVL1392 baculovirus
 CC vector and expression in Sf9 cells in the presence of reduced
 CC glutathione and dithiothreitol.
 XX
 SQ Sequence 72 AA:
 Query Match 100.0%; Score 25; DB 16; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESYT 5
 DB 24 LESYT 27
 RESULT 7
 AAAY9031
 ID AAAY9031 standard; protein; 76 AA.
 XX
 AAAY9031:
 30-MAY-2000 (first entry)
 OS Homo sapiens.
 XX
 DE Amino acid sequence of chemokine receptor ligand MCP-2.
 XX
 KW Chemokine receptor; ligand; inflammatory response; immune effector cell;
 KW secondary tissue damage; central nervous system injury; MCP-2;
 KW CNS inflammatory disease; neurodegenerative disorder; heart disease;
 KW inflammatory eye disease; inflammatory bowel disease;
 KW inflammatory joint disease; inflammatory kidney; renal disease;
 KW inflammatory lung disease; inflammatory nasal disease;
 KW inflammatory thyroid disease; thyroiditis; cytokine-regulated cancer.
 XX
 OS
 XX
 OS Homo sapiens.
 XX
 PN W0200004926-A2.
 PD 04-FEB-2000.
 XX
 PF 21-JUL-1999: 99WO-CA00659.
 XX
 PR 22-JUL-1998: 98US-0120523.
 XX
 PA (USFE-) OSPREY PHARM LTD.
 XX

PI McDonald JR, Coggin PJ;
 XX WPI: 2000-182542/16.
 DR
 XX A new therapeutic agent comprising a conjugate for treating secondary
 PT tissue damage and other disease conditions like Alzheimer's disease,
 PI stroke, Parkinson's disease and atherosclerosis
 XX
 XX Disclosure: Page 60; 204pp; English.
 PS
 XX The present sequence represents a chemokine receptor ligand. The present
 CC ligand can be incorporated into the conjugates of the invention. The
 CC conjugate describes a conjugate, comprising a targeted agent and a
 CC chemokine receptor ligand. The conjugate binds to a chemokine receptor
 CC resulting in internalisation of the targeted agent in cells bearing the
 CC receptor. The conjugates are used for formulating a medicament or for
 CC treating disorders associated with inflammatory responses resulting from
 CC activation, proliferation and migration of immune effector cells. The
 CC disorders or disease states comprise secondary tissue damage such as
 CC central nervous system (CNS) injury, CNS inflammatory eye diseases,
 CC neurodegenerative disorders, heart disease, inflammatory joint diseases,
 CC inflammatory bowel diseases, inflammatory lung diseases, inflammatory
 CC kidney or renal diseases, inflammatory thyroid disease such as thyroiditis, or
 CC cytokine-regulated cancers.
 XX
 SQ Sequence 76 AA:
 Query Match 100.0%; Score 25; DB 21; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESYT 5
 DB 25 LESYT 29
 RESULT 8
 AA021099
 ID AA021099 standard; protein; 76 AA.
 XX
 AC AA021099;
 XX
 DT 19-JUL-2002 (first entry),
 XX
 DE Protein of WT-MCP-2.
 DE
 XX RANTES; neuroprotective; anti-allergic; anti-inflammatory; anti-HIV; human;
 KW chemokine mutant; cationic site; multiple sclerosis; HIV infection;
 KW inflammatory disease; demyelinating disease; allergic.
 XX
 OS Homo sapiens.
 XX
 PN W0200228419-A2.
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001: 2001WO-EP11426.
 XX
 PR 04-OCT-2000: 2000EP-0121065.
 XX
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 XX
 PI Proudfoot A, Wells TNC, Kosco-Villbois M;
 XX WPI: 2002-340073/37.
 DR
 XX A mutant of the human CC chemokine RANTES with two mutations in the
 PT cationic site of the 40% loop is used for treatment of multiple
 PI sclerosis and/or demyelinating diseases -
 XX
 PS Examples: Page 46; 46pp; English.
 XX

CC The invention relates to a truncated and mutated human RANTES (a CC
 CC chemokine mutant), comprising the amino sequence of 91 amino acids as
 CC given in the specification. The CC chemokine mutant RANTES, with two
 CC mutations in the cationic site is useful for the preparation of a
 CC pharmaceutical composition used in treating multiple sclerosis or other
 CC demyelinating diseases. The mutant with single mutations at cationic
 CC sites is used for the treatment of HIV infection and/or other allergic or
 CC inflammatory diseases. This sequence represents an example of a wild-type
 CC chemokine protein used in the invention.

XX Sequence 76 AA:

QY 1 LESYT 5
 DB 25 LESYT 29

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AA021109
 ID AAO21109 standard; Protein; 76 AA.
 AC AAO21109;
 XX
 DT 19-JUL-2002 (first entry)
 DE MCP-2 chemokine protein containing 40's loop and cationic site
 XX
 KW RANTES: neuroprotective; anti-allergic; anti-inflammatory; anti-HIV; human;
 KW chemokine mutant; cationic site; multiple sclerosis; HIV infection;
 KW inflammatory disease; demyelinating disease; allergic
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 36..52
 FT /label= 40's_loop
 FT Misc-difference 46..49
 FT /label= Cationic_site
 FT
 PN WO200228419-A2.
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-EP11428.
 VV
 VV 04-OCT-2000; 2000EP-0141555.
 AX
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 XX
 PI Proudfoot A, Wells TNC, Kosco-Vilbois M;
 DR WPI: 2002-340073/37.
 XX
 XX A mutant of the human CC chemokine RANTES with two mutations in the
 PT cationic site of the 40's loop is used for treatment of multiple
 PT sclerosis and/or demyelinating diseases -
 PI
 PS Disclosure: Fig 1; 46pp; English.

CC The invention relates to a truncated and mutated human RANTES (a CC
 CC chemokine mutant), comprising the amino sequence of 91 amino acids as
 CC given in the specification. The CC chemokine mutant RANTES, with two
 CC mutations in the cationic site is useful for the preparation of a
 CC pharmaceutical composition used in treating multiple sclerosis or other
 CC demyelinating diseases. The mutant with single mutations at cationic
 CC sites is used for the treatment of HIV infection and/or other allergic or
 CC inflammatory diseases. This sequence represents an example of a CC
 CC chemokine protein containing a 40's loop and cationic site of the
 CC invention.

XX Sequence 76 AA:

QY 1 LESYT 5
 DB 25 LESYT 29

RESULT 10
 AAG78390
 ID AAG78390 standard; Protein; 76 AA.
 XX
 AC AAG78390;
 XX
 DT 30-MAY-2002 (first entry)
 DE Human chemokine hMCP-2.
 XX
 KW Antigen presenting cell; APC; chemotaxis; immune response; antigen;
 KW chemokine; antitumor; cytostatic; antimicrobial; toxin;
 KW pathogenic agent; cancer; melanoma; thyroid carcinoma;
 KW renal cell carcinoma; tumor-associated antigen; human; hMCP-2.
 XX
 OS Homo sapiens.
 XX
 PN WO200180882-A2.
 XX
 PD 01-NOV-2001.
 XX
 FE 20-APR-2001; 2001WO-US21000.
 XX
 PR 21-APR-2000; 2000US-1988 (wp).
 PR 12-APR-2001; 2001US-0834814.
 XX
 FA (CHEN-) CHPM/CENTRIX INC.
 PI Schall TJ, Talbot D;
 XX
 WP: 2002-034405/04.
 XX
 ES Claim 7; Fig 2; 60pp; English.

CC The present sequence is that of a human chemokine polypeptide designated
 CC hMCP-2, used in the compositions of the invention. The specification
 CC describes the use of a composition containing an antigen presenting cell
 CC (APC)-chemotaxis for inducing an immune response to an antigen in a
 CC subject, in the manufacture of a medicament, where the APC-chemotaxis is
 CC a chemokine polypeptide or its variant, or a polynucleotide encoding the
 CC chemokine polypeptide or its variant. The invention has antitumor,
 CC cytostatic and antimicrobial activity. The compositions of the invention
 CC can contain one or more antigens for antigen containing polynucleotides
 CC and may be administered in the same mixture as the APC-chemotaxis or
 CC separately. Synthetic (minimally) or naturally occurring chemokines may be
 CC used. The polypeptides of the invention are used to induce an immune
 CC response by recruiting APCs to areas of antigen contact. The compositions
 CC of the invention are useful for providing protection from foreign
 CC infectious pathogenic agents (bacteria, virus, etc.) prior to expected
 CC or possible exposure; for treating cancers, melanomas,
 CC thyroid carcinomas, lung and breast cancers, renal cell carcinomas etc.,
 CC here the antigen present in the composition is a tumor associated
 CC antigen. The use of APC-chemotaxis provides an accelerated immune
 CC response in a host following administration of antigen, a more effective
 CC response to administration of exposure to very small quantities of
 CC an antigen (e.g., toxin or pathogen) due to increased antigen uptake by

CC APCs, and more effective anti-tumour therapies.
 XX
 SQ Sequence 76 AA:

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LE5YT 5
 11111
 DB 25 LE5YT 29

RESULT 11
 AAG68353
 ID AAG68353 standard; peptide: 76 AA
 XX
 AC AAG68353:
 VX 01-MAR-2002 (first entry)

Human monocyte chemotactic protein 2 (MCP2) amino acid sequence.

XX Human: mouse; monocyte chemotactic protein 2; (MCP2; MCP2; m310; mMEC;
 XX chimeric chemokine; immune response; cytostatic; antitubercill;
 XX immunostimulant; vaccine; cancer; breast cancer; lung cancer;
 XX carcinoma; melanoma; tumour; hyaluronin.

XX Homo sapiens.

XX WO2001060847-A2.

XX 01-NOV-2001.

XX 12-APR-2001; 2001WD-0512162.

XX 21-APR-2000; 2000US-198839P.

XX (CHEM-) CHEMOCENTRIX INC.

XX Schall TJ, Talbot D;

XX WPI: 2002-049246/06.

XX Inducing an immune response to an antigen to confer non-specific

XX protection while the body is generating the adaptive response;

XX comprises administering a composition containing an antigen-presenting

XX cell chemotaxin

XX Example 5; Fig 2; 60pp; English.

XX The present invention describes a method for inducing an immune response
 XX to an antigen in a subject. The method comprises administering a
 XX composition containing an antigen-presenting cell chemotaxin

XX (APC-chemotaxin), where the APC-chemotaxin is a chemokine polypeptide,
 XX its variant, or a polynucleotide encoding the chemokine polypeptide or

XX its variant. The APC-chemotaxin has cytostatic, antimicrobial and
 XX immunostimulant activities, and can be used in vaccine production.

XX The method can be used for inducing or enhancing an immune response,
 XX or for providing protection from exogenous foreign infectious pathogenic

XX agents prior to expected or possible exposure. The method may be used to treat cancers
 XX displaying symptoms of exposure. The method may be used to treat cancers

XX e.g. breast cancer, lung cancer, carcinomas, melanomas, and tumours.

XX The present sequence represents a human monocyte chemotactic protein 2
 XX (MCP2) amino acid sequence, which is used in an example from the

XX present invention for the design of hybridomas (chimeric chemokines).

XX Sequence 76 AA:

Query Match 100.0%; Score 25; DB 23; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LE5YT 5
 11111
 DB 25 LE5YT 29

RESULT 12
 AAB15786
 ID AAB15786 standard; Protein 77 AA.

XX AAB15786:

XX 17-JAN-2001 (first entry)

XX Human chemokine MCP-2 SEQ. ID NO: 17.

XX Macrophage recruitment; chemokine derivative; MCP-1; osteoporosis;
 XX monocyte chemotactic protein-1; inflammation; atherosclerosis; HIV;
 XX AIDS; stroke; psoriasis; autoimmune disease; hypertension; endotoxaemia;
 XX basophil-mediated disease; myocardial infarction; acute ischaemia;
 XX rheumatoid arthritis; coagulation.

XX Homo sapiens.

XX Key location/qualifiers

XX Misc-difference 47 /note="encoded by CAA"

XX WO200042071-A2.

XX 20-JUL-2000.

XX 12-JAN-2000; 2000WD-0500831.

XX 12-JAN-1999; 99US-0229071.

XX 17-MAR-1999; 99US-0271102.

XX 01-DEC-1999; 99US-0452416.

XX (NEOR-) NEORX CORP.

XX Grainger DJ, Tatalick LM;

XX WPI: 2000-499101/44.

XX N-PSDB: AAA74886.

XX New peptide 3, amide and heterocyclic compounds and saccharide

XX conjugates used for inhibiting chemokine induced activity and for

XX treating e.g. stroke, vascular diseases, autoimmune diseases and tumour

XX growth

XX Example 1; Page 134; 387pp; English.

XX The present invention concerns the identification of a number of
 XX chemokines which can be used to produce derivatives, agonists and

XX antagonists which are then useful in disease treatment. The chemokines
 XX include sequences AAB15786-B15794, AAB15803-B15813 and AAB15831-B15845.

XX These chemokine derivatives can be used to treat diseases such as
 XX autoimmune diseases, atherosclerosis, osteoporosis, HIV infection and

XX AIDS, psoriasis, inflammatory diseases, hypertension, basophil-mediated
 XX diseases, endotoxaemia, myocardial infarction, acute ischaemia and

XX rheumatoid arthritis, and can be used to prevent strokes and as
 XX contraceptive. The coding sequences for the chemokines can be used in

XX gene therapy for the same diseases, as well as in the production of
 XX animal models.

XX Sequence 77 AA:

Query Match 100.0%; Score 25; DB 23; Length 77;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
 ID ABB80903
 AC ABB80903 standard: Protein: 77 AA.
 XX
 XX ABB80903:
 DI 08-UCI-4002 (first entry)
 XX
 DE Human MCP2 fragment.
 XX
 KM Eotaxin; eosinophil; chemotaxis; cytostatic; antiinflammatory; cardiact;
 KM antiallergic; immunosuppressive; antiallergic; antiparasitic; histamine;
 KM dermatologic; vasotropic; monocyte chemoattractive protein; MCP;
 KM gene therapy; antianaphylactic.
 XX
 OS Homo sapiens.
 XX
 PN US6403782-B1.
 XX
 Y 11-JUN-2002.
 XX
 FE 04-AUG-1999: 99US-0366887.
 XX
 PR 22-MAY-1995: 95US-000449P.
 PR 01-SEP-1995: 95US-0522713.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (GERH) GEN HOSPITAL CORP.
 XX
 PI Luster AD, Leder P, Rothenberg M, Garcia E;
 XX
 DP WPI: 2002-565447/60.
 XX
 PT New DNA encoding murine, guinea pig or human eotaxin polypeptides,
 PT useful for treating inflammation and tumorigenesis and in anticancer
 PT gene therapy.
 XX
 PS Disclosure: Fig 1B; 42pp; English.
 XX
 CC The invention relates to polynucleotides encoding murine, guinea pig or
 CC human eotaxin polypeptides. The eotaxin polynucleotides are useful for
 CC modulating eosinophil chemotaxis, for increasing eosinophil chemotactic
 CC events, and for improving prognosis with tumors in patients. They are
 CC also useful for treating inflammation and tumorigenesis, and for
 CC reducing inflammation and cytotoxic damage caused by eosinophils, for
 CC e.g. during asthmatic reactions, eosinophilic pneumonia and allergic
 CC diseases, inflammatory bowel diseases, atopic dermatitis, urticaria,
 CC vasculitis, parasitic infections and eosinophil cardiac diseases. The
 CC eotaxin polynucleotides are also useful for modulating histamine release
 CC by modulating eotaxin activity or expression during anaphylaxis,
 CC urticaria and allergic reactions. They are useful for detecting and
 CC monitoring eosinophil mediating conditions, and in anti-cancer gene
 CC therapy. The present sequence represents a human monocyte chemotactic
 CC protein (MCP)2 fragment.
 CC
 XX
 SQ Sequence 77 AA:
 XX
 XX Query Match 100.0%; Score 25; DB 23; Length 77,
 XX Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 LESYT 5
 XX | | | | |
 DB 26 LESYT 30
 XX
 RESULT 14
 ID AAG78395 standard: protein: 77 AA.
 XX
 AC AAG78395:
 XX
 AC AAG78395:

XX
 DT 30-MAY-2002 (first entry)
 XX
 DE Mouse/human hybridoma mAb/hMCP-2.
 XX
 KM Antigen presenting cell; APC; chemotaxis; immune response; antigen;
 KM chemokine, antitumor; cytostatic; antimicrobial; toxic;
 KM pathogenic agent; cancer; melanoma; thyroid carcinoma; hybridoma;
 KM renal cell carcinoma; tumour-associated antigen; human; mouse.
 XX
 OS ChimERIC - Homo sapiens.
 OS ChimERIC - Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Region 1..13
 FT /note= "Corresponds to residues 1-13 of mAb
 FT (AAG78392)"
 FT
 FT Region 12..77
 FT /note= "Corresponds to residues 11-76 of hMCP2
 FT (AAG78390)"
 XX
 PN W020100832-A2.
 XX
 PD 01-NOV-2001.
 XX
 FE 26-APR-2001: 2001W-0512950.
 XX
 PR 21-APR-2000: 2000US-198849P.
 PR 12-APR-2001: 2001US-083484.
 XX
 PA (CHEM-) CHEMOCENTRIPYX INC
 XX
 PI Schall TJ, Talbot D;
 XX
 DR WPI: 2002-034405/04.
 XX
 PT Use of a composition containing an antigen presenting cell chemotaxin
 PT which is a chemokine polypeptide or its variant or a polynucleotide
 PT encoding the protein, for inducing immune response to an antigen in a
 PT subject.
 XX
 PS Example 5; Fig 2; 60pp; English.
 XX
 CC The present sequence is that of a chimERIC chemokine polypeptide
 CC (hybrid) derived from human chemokine hMCP-2 (AAG78390) and murine
 CC chemokine mAb (AAG78392) used in the compositions of the invention. The
 CC specification describes the use of a composition containing an antigen
 CC presenting cell (APC)-chemotaxin for inducing an immune response to an
 CC antigen in a subject, in the manufacture of a medicament, where the
 CC APC-chemotaxin is a chemokine polypeptide or its variant, or a
 CC polynucleotide encoding the chemokine polypeptide or its variant. The
 CC invention has antitumor, cytostatic and antimicrobial activity. The
 CC compositions of the invention can contain one or more antigens (or
 CC antigen containing polynucleotides) and may be administered in the same
 CC mixture as the APC-chemotaxin or separately. Synthetic (minetic) or
 CC naturally occurring chemokines may be used. The polypeptides of the
 CC invention are used to induce an immune response by recruiting APCs to
 CC areas of antigen contact. The compositions of the invention are useful
 CC for providing protection from foreign infectious pathogenic agents
 CC (bacteria, virus, etc.) prior to expected or possible exposure; for
 CC treating cancers, melanomas, thyroid carcinomas, lung and breast cancers,
 CC renal cell carcinomas etc., here the antigen present in the composition
 CC is a tumour-associated antigen. The use of APC-chemotaxin provides an
 CC accelerated immune response in a host following administration of
 CC antigen, a more effective response to administration of or exposure to,
 CC very small quantities of an antigen (e.g., toxin or pathogen) due to
 CC increased antigen uptake by APCs, and more effective anti-tumour
 CC therapies.
 CC
 XX
 SQ Sequence 77 AA:
 XX
 XX Query Match 100.0%; Score 25; DB 23; Length 77;
 XX Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 XX

Search completed: April 8, 2003, 12:01:37
 Job time: 45.333 secs

Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LESYT 5
 11111
 Db 26 LESYT 30

RESULT 15

AAG68358

ID AAG68358 standard: peptide; 77 AA.

XX AAG68358:

DT 01-MAR-2002 (first entry)

DE chimeric chemokine mMDC/hMCP2 amino acid sequence.

XX Human: mouse; monocytic chemotactic protein 2; hMCP2; MCP2; MCI0; mMDC;

KW chimeric chemokine; immune response; cytostatic; antimicrobial;

KW immunostimulant; vaccine; cancer; breast cancer; lung cancer;

KW carcinoma; melanoma; tumour; hybridoma.

XX Chimeric - Mus sp.

OS Chimeric - Homo sapiens.

XX WO200180887-A2.

PD 01-NOV-2001.

XX 12-APR-2001; 2001WO-US12162.

XX 21-APR-2000; 2000US-198939P.

XX (CHEM-) CHEMOCENTRIX INC.

XX Schall TJ. Talbot D.

XX WPI: 2002-043246/06.

XX Inducing an immune response to an antigen to confer non-specific

PT protection while the body is generating the adaptive response,

PT comprises administering a composition containing an antigen-presenting

PT cell chemotaxin -

XX Example 5: Fig 2; 60pp; English.

XX The present invention describes a method for inducing an immune response

XX to an antigen in a subject. The method comprises administering a

XX composition containing an antigen-presenting cell chemotaxin

XX (APC-chemotaxin), where the APC-chemotaxin is a chemokine polypeptide,

XX its variant, or a polynucleotide encoding the chemokine polypeptide or

XX its variant. The APC-chemotaxin has cytostatic, antimicrobial and

XX immunostimulant activities, and can be used in vaccine production.

XX The method can be used for inducing or enhancing an immune response,

XX or for providing protection from exogenous exposure, or to individuals

XX agents prior to expected or possible exposure. The method may be used to treat cancers

XX displaying symptoms of exposure. The method may be used to treat cancers

XX e.g. breast cancer, lung cancer, carcinomas, melanomas, and tumours.

XX The present sequence represents an exemplary chimeric chemokine

XX designated mMDC/hMCP2, which is given in an example from the present

XX invention for the design of hybridomas (chimeric chemokines).

SQ Sequence 77 AA:

Query Match 100.0%; Score 25; DB 23; Length 77;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
11111
Db 26 LESYT 30

GenCorp Version 5.1.4-PS-4SPD
Copyright (c) 1993 - 2003 Computer Ltd.

CM protein - protein search, using SW model

Run on: April 8 2003, 11:53:45 : Search time 15.467 seconds
(without alignments)
9.543 Million cell updates/sec

Title: US-09-647-749A-1

Perfect score: 25

Sequence: 1 LRSYT S

Scoring table: RUSOM62

Gapop 10.0 : Gapext 0.5

Searched: 262574 seqs, 29422222 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 400000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents, AA:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | FP | ID | Description |
|------------|-------|--------------------|------|----|---------------------|
| 1 | 25 | 100.0 | 74 | 2 | US-08-615-232A-6 |
| 2 | 25 | 100.0 | 74 | 3 | US-08-470-423-6 |
| 3 | 25 | 100.0 | 74 | 1 | US-08-480-449-20 |
| 4 | 25 | 100.0 | 76 | 2 | US-08-716-188-3 |
| 5 | 25 | 100.0 | 76 | 2 | US-08-660-542-20 |
| 6 | 25 | 100.0 | 76 | 4 | US-08-470-603-20 |
| 7 | 25 | 100.0 | 77 | 1 | US-08-347-4928-9 |
| 8 | 25 | 100.0 | 77 | 2 | US-08-421-144A-6 |
| 9 | 25 | 100.0 | 77 | 2 | US-08-798-143-9 |
| 10 | 25 | 100.0 | 77 | 4 | US-08-366-882A-12 |
| 11 | 25 | 100.0 | 255 | 4 | US-09-327-857-488 |
| 12 | 25 | 100.0 | 634 | 4 | US-08-951-871-6 |
| 13 | 25 | 100.0 | 634 | 4 | US-08-134-001C-4632 |
| 14 | 25 | 100.0 | 2499 | 4 | US-09-562-737-116 |
| 15 | 25 | 92.0 | 146 | 4 | US-08-781-562-4 |
| 16 | 25 | 92.0 | 280 | 1 | US-09-134-001C-3736 |
| 17 | 25 | 92.0 | 316 | 4 | US-08-436-664-20 |
| 18 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 19 | 25 | 92.0 | 876 | 2 | US-08-436-664-34 |
| 20 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 21 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 22 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 23 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 24 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 25 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 26 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |
| 27 | 25 | 92.0 | 876 | 2 | US-08-436-664-32 |

| | | | | | |
|----|----|------|-----|---|-------------------|
| 28 | 25 | 92.0 | 876 | 5 | PCT-US95-04080-20 |
| 29 | 25 | 92.0 | 876 | 5 | PCT-US95-04080-32 |
| 30 | 25 | 92.0 | 876 | 5 | PCT-US95-04080-34 |
| 31 | 25 | 92.0 | 954 | 1 | US-09-057-989-2 |
| 32 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 33 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 34 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 35 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 36 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 37 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 38 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 39 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 40 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 41 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 42 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 43 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 44 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |
| 45 | 25 | 88.0 | 111 | 1 | US-08-456-424-4 |

ALIGNMENTS

RESULT 1
US-08-615-232A-6
Sequence 6, Application US-08615232A
Patent No. 5993814
GENERAL INFORMATION:
APPLICANT: WILLIAMS, T. MOTHY J.
APPLICANT: JOSE, PETER J.
APPLICANT: GRIFFITHS-JOHNSON, DAVID A.
APPLICANT: HEYMAN, JOHN I.
TITLE OF INVENTION: CHEMOTACTIC CYTOKINE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ALBRECHT, RIXON & VANCE, P. C.
STREET: 1100 NORTH GENE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (HW)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08 931894
FILING DATE: 14-SEP-1993
APPLICATION NUMBER: 08 9408602
FILING DATE: 29-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 02955
REFERENCE/BOOK NUMBER: 550-32
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-615-232A-6
Query Match: 100.0% Score 25; DB 2; Length 74;
Best Local Similarity: 100.0% Pred. No. 40;

Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 LESYT 5
|||||
DB 23 LESYT 27

RESULT 2

US-08-470-323-6
Sequence 6, Application US/08470323A
Patent No. 6031080
GENERAL INFORMATION:
APPLICANT: WILLIAMS, TIMOTHY J.
APPLICANT: JOSE, PETER J.
APPLICANT: GRIFFITHS-JOHNSON, DAVID A.
APPLICANT: HSUAN, JOHN J.
TITLE OF INVENTION: CHEMOTACTIC CYTOKINE
FILE REFERENCE: 550-33
CURRENT APPLICATION NUMBER: US-08-470-323A
CURRENT FILING DATE: 1995-06-06
EARLIER APPLICATION NUMBER: PCT/US94/02006
EARLIER FILING DATE: 1994-09-14
EARLIER APPLICATION NUMBER: GB 9318984.3
EARLIER FILING DATE: 1993-09-14
EARLIER APPLICATION NUMBER: GB 9406699.2
EARLIER FILING DATE: 1994-04-29
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 74
TYPE: PRT
ORGANISM: human
US-08-470-323-6

Query Match 100.0%, Score 25; DB 3; Length 74;
Best Local Similarity 100.0%, Pred. No. 40;
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 1 LESYT 5
|||||
DB 23 LESYT 27

RESULT 3

US-08-480-449-20
Sequence 26, Application US/08480449
Patent No. 5686927
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,449
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 278/6,72779
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: "Hu MCP-2"

US-08-480-449-20

Query Match 100.0%, Score 25; DB 1; Length 76;
Best Local Similarity 100.0%, Pred. No. 41;
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 1 LESYT 5
|||||
DB 25 LESYT 29

RESULT 4

US-08-716-188-3
Sequence 3, Application US/08716188
Patent No. 5908829
GENERAL INFORMATION:
APPLICANT: KELLY, RODNEY W.
TITLE OF INVENTION: USE OF MCP-1 FOR INDUCING RIPENING OF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANIERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/716,188
FILING DATE: 30-SEP-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00733
FILING DATE: 31-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: G0 9406463.1
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 46,663
REFERENCE/DOCKET NUMBER: 117-219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-716-188-3

Query Match 100.0%, Score 25; DB 2; Length 76;
Best Local Similarity 100.0%, Pred. No. 41;
Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 1 LESYT 5
|||||
Db 25 LESYT 29

RESULT 5
US-08-660-542-20
Sequence 20, Application US/08060042
Patent No. 5932703
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE AND CHEMOKINE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 213 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/660,542
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/558,658
FILING DATE: 16-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,620
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 27466,44219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc:feature
OTHER INFORMATION: "Hu MCP-2"
US-08-660-542-20
Query Match 100.0%; Score 25; DB 2; Length 76,
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LESYT 5
|||||
Db 25 LESYT 29
RESULT 6
US-08-479-603-20
Sequence 20, Application US/08479603
Patent No. 6320023
GENERAL INFORMATION:
APPLICANT: Godiska, Ronald

APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 213 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,603
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 27466/32780
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc:feature
OTHER INFORMATION: "Hu MCP 2"
US-08-479-603-20
Query Match 100.0%; Score 25; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LESYT 5
|||||
Db 25 LESYT 29
RESULT 7
US-08-347-4928-9
Sequence 9, Application US/08447492B
Patent No. 5602008
GENERAL INFORMATION:
APPLICANT: Wilde, Craig J.
APPLICANT: Hawkins, Phillip F.
APPLICANT: Bandman, Olga
APPLICANT: Sellman, Jeffrey J.
TITLE OF INVENTION: EXPRESSED CHEMOKINES, THEIR
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:

APPLICATION NUMBER: 05/08/747 4928
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA: 05/03/241
APPLICATION NUMBER: 07-SEP-1994
FILING DATE: 08/320.011
APPLICATION NUMBER: 08/320.011
FILING DATE: 05-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PF-0024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: GI 126829
US-08-447-4928-9
Query Match
Best Local Similarity 100.0%; Score 25; DB 1; Length 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
CY 1 LEAST 5
DB 26 LEAST 30
RESULT 8
US-08-421-144A-6
Sequence 6, Application US/08421144A
Patent No. 5874211
GENERAL INFORMATION:
APPLICANT: BANDMAN, OLGA
APPLICANT: COLEMAN, ROGER
APPLICANT: STUART, SUSAN G.
TITLE OF INVENTION: NEW CHEMOKINE EXPRESSED IN EOSINOPHILS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,144A
FILING DATE: 13-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0031 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-421-144A-6
Query Match
Best Local Similarity 100.0%; Score 25; DB 2; Length 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
CY 1 LEAST 5
DB 26 LEAST 30
RESULT 9
US-08-798-143-9
Sequence 9, Application US/08798143
Patent No. 5936068
GENERAL INFORMATION:
APPLICANT: Wilde, Craig J.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Bandman, Olga
APPLICANT: Seilhamer, Jeffrey J.
TITLE OF INVENTION: EXPRESSED CHEMOKINES, THEIR
TITLE OF INVENTION: PROMOTION AND USES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/798,143
FILING DATE: 10-FEB-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/447,492
FILING DATE: 29-NOV-1994
APPLICATION NUMBER: 08/403,241
FILING DATE: 07-SEP-1994
APPLICATION NUMBER: 08/420,011
FILING DATE: 05-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: PF-0024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: GI 126829
US-08-798-143-9
Query Match
Best Local Similarity 100.0%; Score 25; DB 2; Length 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
CY 1 LEAST 5

```

Db      26 LE5YT 30

RESULT 10
US-09-366-887A-12
Sequence 12, Application US/09366887A
Patent No. 6403782
GENERAL INFORMATION:
APPLICANT: LUSTER, ANDREW D.
APPLICANT: LEDER, PHILIP
APPLICANT: POTENBERG, MARC
APPLICANT: GARCIA, EDUARDO
TITLE OF INVENTION: EDAXIN: AN FOSFONICIL CHEMOTHERAPY
FILE REFERENCE: 00483/025002
CURRENT APPLICATION NUMBER: US/09/366-887A
PRIOR FILING DATE: 1999-08-04
PRIOR APPLICATION NUMBER: 60/090,444
PRIOR FILING DATE: 1995-06-22
PRIOR APPLICATION NUMBER: 08/522,713
PRIOR FILING DATE: 1995-09-01
PRIOR APPLICATION NUMBER: 08/522,714
PRIOR FILING DATE: 1998-06-16
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-09-366-887A-12

Query Match      100.0% Score 25: DB 4: Length 77:
Best local similarity 100.0%, Freq. No. 42:
Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY      1 LE5YT 5
        |||||
Db      26 LE5YT 30

RESULT 11
US-09-227-357-488
Sequence 488, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PST/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,940
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08

```

```

EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-19
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,795
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: Patentia Ver. 2.0
SEQ ID NO 488
LENGTH: 255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-227-357-488

Query Match      100.0% Score 25: DB 4: Length 255:
Best local similarity 100.0%, Freq. No. 1,440:
Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0:

QY      1 LE5YT 5
        |||||
Db      163 LE5YT 167

RESULT 12
US-08-951-871-6
Sequence 6, Application US/08951871
Patent No. 5866398
GENERAL INFORMATION:
APPLICANT: XIAO, SHUANG-YONG
APPLICANT: XIAO, JIAN-PING
TITLE OF INVENTION: METHOD FOR CLONING
TITLE OF INVENTION: AND PRODUCING THE US11 RESTRICTION ENDONULEASE IN F.
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: New England Biolabs, Inc.
STREET: 32 Tozer Road

```

```

CITY: Beverly
STATE: MA
COUNTRY: US
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951.871
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Gregory D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-131
TELECOMMUNICATION INFORMATION:
TELEPHONE: 978-927-5054
TELEFAX: 978-927-1705
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-951-871-6

Query Match          100.0%  Score 25:  DB 2:  Length 301:
Best Local Similarity 100.0%:  Pred. No. 1.7e+02:
Matches 5:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESYT 5
DB 216 LESYT 220

RESULT 13
US-09-134-001C-4633
Sequence 4633: Application US/09/134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIORITY APPLICATION NUMBER: US 60/7064,964
PRIORITY FILING DATE: 1997-11-08
PRIORITY APPLICATION NUMBER: US 60/7055,779
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4633
LENGTH: 538
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4633

Query Match          100.0%  Score 25:  DB 4:  Length 538:
Best Local Similarity 100.0%:  Pred. No. 3e+02:
Matches 5:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESYT 5
DB 291 LESYT 295

```

```

RESULT 14
5180808-2
Patent No. 5180808
APPLICANT: ROOSLAHNT, EREK I.
TITLE OF INVENTION: VESICULAR CORE PROTEIN, NUCLEIC ACID
SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VESSICAN
ANTIBODIES, AND METHODS OF DETECTING THE SAME
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/441,179
FILING DATE: 27-NOV-1989
SEQ ID NO: 2
LENGTH: 2409
5180808-2

Query Match          100.0%  Score 25:  DB 6:  Length 2409:
Best Local Similarity 100.0%:  Pred. No. 1.4e+03:
Matches 5:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESYT 5
DB 1203 LESYT 1207

RESULT 15
US-09-562-737-116
Sequence 116: Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 116
LENGTH: 185
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-116

Query Match          92.7%  Score 23:  DB 4:  Length 185:
Best Local Similarity 80.1%:  Pred. No. 3e+02:
Matches 4:  Conservative 1:  Mismatches 0:  Indels 0:  Gaps 0:

QY 1 LESYT 5
DB 86 MESYT 90

Search completed: April- 8, 2003, 12:05:51
Job time : 16.4167 secs

```

Genome version 5.1.4.25.4578
Copyright (c) 1993 - 2003 Genengen Ltd.

OM protein - protein search, using sw model

Run on: April 8 2003 12:03:50 Search time 9.10467 seconds
(without alignments)
33,347 Million cell updates/sec

Title: US-09-647-749a-1

Perfect score: 25

Sequence: 1 LESYT 5

Scoring table: BLOSUM62

Gap: 10.0, Offset 0.0

Searched: 248912 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248912

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published Applications_AA:

- 1: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/prodata/1/pubppa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 5: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 9: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 10: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 11: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 12: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
- 14: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*

Pre1 N is the number of results predicted by query to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 25 | 100.0 | 74 | 10 | US-09-647-741-45820 |
| 2 | 25 | 100.0 | 74 | 10 | US-09-647-741-45820 |
| 3 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 4 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 5 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 6 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 7 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 8 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 9 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 10 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 11 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 12 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 13 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 14 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 15 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 16 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 17 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 18 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |
| 19 | 25 | 100.0 | 74 | 9 | US-09-792-793A-21 |

ALIGNMENTS

| | | | | | | |
|----|----|------|-----|----|-------------------|-------------------|
| 20 | 23 | 92.0 | 486 | 10 | US-09-801-468-399 | Sequence 398, App |
| 21 | 23 | 92.0 | 486 | 10 | US-09-801-468-399 | Sequence 10668, A |
| 22 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 40, Appl |
| 23 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 2, Appl1 |
| 24 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 2, Appl1 |
| 25 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 19, Appl |
| 26 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 45457, A |
| 27 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 4, Appl |
| 28 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 45840, A |
| 29 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 10, Appl |
| 30 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 147, App |
| 31 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 708, App |
| 32 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 33 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 34 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 35 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 36 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 37 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 38 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 39 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 40 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 41 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 42 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 43 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 44 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |
| 45 | 23 | 88.0 | 12 | 9 | US-09-945-159A-2 | Sequence 225, App |

PRIOR APPLICATION NUMBER: PRT/US01/000670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-10
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46020
LENGTH: 68
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC00H494.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EST_HUMAN HIT: AA57020.1, EVALUATE 1.00e-07
US-09-647-749a-1

Query Match 100.0% Score 25; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 25 LESYT 29

RESULT 2
US-09-195-457-6
Sequence 6, Application US/09195457
Patent No. US20020081623A1
GENERAL INFORMATION:
APPLICANT: WILLIAMS, TIMOTHY J.
APPLICANT: JOSE, PETER J.
APPLICANT: GRIFFITHS-JOHNSON, DAVID A.
APPLICANT: HSUAN, JOHN J.
TITLE OF INVENTION: CHEMOTACTIC CYTOKINE
FILE REFERENCE: 550-33
CURRENT APPLICATION NUMBER: US/09/195,457
CURRENT FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/470,323
PRIOR FILING DATE: 1995-06-06
PRIOR APPLICATION NUMBER: PRT/US84/02006
PRIOR FILING DATE: 1994-09-14
PRIOR APPLICATION NUMBER: GB 9318984.3
PRIOR FILING DATE: 1993-09-14
PRIOR APPLICATION NUMBER: GB 94086902.2
PRIOR FILING DATE: 1994-04-29
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 6
LENGTH: 74
TYPE: PRT
ORGANISM: human
US-09-195-457-6

Query Match 100.0% Score 25; DB 10; Length 74;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 23 LESYT 27

RESULT 3
US-09-792-793A-21
Sequence 21, Application US/09792793A
Patent No. US20020168170A1
GENERAL INFORMATION:
APPLICANT: McDonald, John R.
APPLICANT: Coggin, Philip

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
FILE REFERENCE: 25020-601D
CURRENT APPLICATION NUMBER: US/09/792,793A
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 76
TYPE: PRT
ORGANISM: homo sapien
FEATURE:
OTHER INFORMATION: Human Chemokine Polypeptide: MCP-2
US-09-792-793A-21

Query Match 100.0% Score 25; DB 9; Length 76;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 25 LESYT 29

RESULT 4
US-08-927-939-17
Sequence 17, Application US/08927939
Patent No. US2001000640A1
GENERAL INFORMATION:
APPLICANT: Granger, David L.
APPLICANT: Tatalick, Lauren Marie
TITLE OF INVENTION: Compounds and methods to inhibit or
TITLE OF INVENTION: augment an inflammatory response.
FILE REFERENCE: 295-0220S1
CURRENT APPLICATION NUMBER: US/08/927,939
CURRENT FILING DATE: 1997-09-11
NUMBER OF SEQ ID NOS: 83
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-08-927-939-17

Query Match 100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 5
US-10-057-275-10
Sequence 10, Application US/10057275
Patent No. US2002015545A1
GENERAL INFORMATION:
APPLICANT: Coleman, Roger
APPLICANT: Bandman, Olga
APPLICANT: Wilde, Craig G.
TITLE OF INVENTION: NEW CHEMOKINES EXPRESSED IN PANCREAS
NUMBER OF SEQUENCE ADDRESSES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,275
FILING DATE: 25-Jan-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/390,740A
FILING DATE: February 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: FF-0027 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: MCP-2
SEQUENCE DESCRIPTION: SEQ ID NO: 10.

Query Match 100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 6
US-09-834-794A-24
Sequence 24, Application US/09/044,794A
Publication No. US2003026777A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US4
CURRENT APPLICATION NUMBER: US/09/044,794A
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-09-834-794A-24

Query Match 100.0% Score 25; LB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 7

US-09-920-137A-10
Sequence 10, Application US/09/020137A
Publication No. US2003004977A1
GENERAL INFORMATION:
APPLICANT: Coleman, Roger
APPLICANT: Bandman, Olga
APPLICANT: Wilde, Craig
TITLE OF INVENTION: NEW THROMBINES EXPRESSED IN HUMAN PNAS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Icyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,137A
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954
REFERENCE/DOCKET NUMBER: FF-0027 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: MCP-2
US-09-920-137A-10

Query Match 100.0% Score 25; DB 9; Length 77;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
DB 26 LESYT 30

RESULT 8
US-09-834-795A-24
Sequence 24, Application US/09/044,795A
Patent No. US20020076710A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US4
CURRENT APPLICATION NUMBER: US/09/044,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24

us-09-647-749a-1.rapb

```

:SOFT 14
:US-09-764-868-875
: Sequence 875, Application US/09764868
: Patent No. US2002016871A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Rosen et al.
:
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
:
: FILE REFERENCE: PT232
:
: CURRENT APPLICATION NUMBER: US/09/764,868
:
: CURRENT FILING DATE: 2001-01-17
:
: Prior application data removed - refer to PALM or title wrapper
:
: NUMBER OF SEQ. ID NOS: 1510
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ. ID NO 875
:
: LENGTH: 151
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: SITE
: LOCATION: (117)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
:
: NAME/KEY: SITE
: LOCATION: (127)
:
: OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
:
: US-09-764-868-875

```

```

RESULT 12
US-09-895-913A-184
; Sequence 184, Application US/09895913A
; Patent No. US20020160456A1
;
GENERAL INFORMATION:
;
APPLICANT: Kleantous, Harold
;
APPLICANT: Al-Garawi, Amal
;
APPLICANT: Miller, Charles
;
APPLICANT: Tomb, Jean Francois
;
APPLICANT: Oomen, Raymond P.
;
TITLE OF INVENTION: Identification of Polynucleotides

```

```

1 TITLE OF INVENTION: Encoding No. US2002019760A1 Helicobacter Pylorides in the
2 FILE REFERENCE: 0514/045002
3 CURRENT APPLICATION NUMBER: US/09/895,913A
4 PRIOR FILING DATE: 2001-06-29
5 PRIOR APPLICATION NUMBER: US 08/781,227
6 PRIOR FILING DATE: 1997-06-24
7 NUMBER OF SEQ ID NOS: 368
8 SOFTWARE: FastSeq for Windows Version 4.0
9 SEQ ID NO: 184
10 LENGTH: 249
11 TYPE: PRT
12 ORGANISM: Helicobacter pylori
13 US-09-895-913A-184

```

```

Query Match      100.0% Score 25; DB 9; Length 249.
Best Local Similarity 100.0% Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 LEAST 5
11111
b 204 LEAST 208

```

```

RESULT 13
US-09-983-802-488
1 Sequence 488 Application US/0983802
2 Publication No. US2002019760A1
3 GENERAL INFORMATION:
4 APPLICANT: Fischer et al.
5 TITLE OF INVENTION: 123 Human Secreted Proteins
6 FILE REFERENCE: P210P1
7 CURRENT APPLICATION NUMBER: US-09/983,802
8 PRIOR FILING DATE: 2001-10-25
9 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/227,357
10 PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-08
11 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13684
12 PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
13 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
14 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
15 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,793
16 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
17 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
18 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
19 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
20 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
21 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
22 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
23 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
24 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
25 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
26 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
27 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
28 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
29 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
30 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
31 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,923
32 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
33 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931
34 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
35 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,932
36 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
37 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
38 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
39 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
40 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
41 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
42 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
43 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
44 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
45 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,923
46 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
47 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
48 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
49 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
50 PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
51 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
52 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
53 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723

```

```

1 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
2 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
3 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
4 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
5 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
6 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953
7 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
8 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
9 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
10 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
11 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
12 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
13 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
14 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,360
15 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
16 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,884
17 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
18 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
19 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
20 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,374
21 PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
22 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,664
23 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
24 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,785
25 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
26 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,660
27 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
28 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,661
29 PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
30 SOFTWARE: PatentIn Ver. 2.1
31 SEQ ID NO: 488
32 LENGTH: 255
33 TYPE: PRT
34 ORGANISM: Homo sapiens
35 US-09-983-802-488

```

```

Query Match      100.0% Score 25; DB 9; Length 255.
Best Local Similarity 100.0% Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 LEAST 5
11111
Db 163 LEAST 167

```

```

RESULT 14
US-09-738-626-6675
1 Sequence 6675 Application US/09738675
2 Publication No. US2002019760A1
3 GENERAL INFORMATION:
4 APPLICANT: NAKAYAMA, SATOSHI
5 APPLICANT: MIZOGUCHI, HIROSHI
6 APPLICANT: ANDO, SEIKO
7 APPLICANT: HAYASHI, MIKIKO
8 APPLICANT: OCHIDA, KEIKO
9 APPLICANT: YOKOI, HARUHIKO
10 APPLICANT: TATEISHI, NAOKO
11 APPLICANT: SENOH, AKIHITO
12 APPLICANT: IKEDA, MASATO
13 APPLICANT: OZAKI, AKIO
14 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
15 FILE REFERENCE: 249-125
16 CURRENT APPLICATION NUMBER: 09/738,675
17 PRIOR FILING DATE: 2000-12-18
18 PRIOR APPLICATION NUMBER: JP 99/377484
19 PRIOR FILING DATE: 1999-12-18
20 PRIOR APPLICATION NUMBER: 31 80/159162
21 PRIOR FILING DATE: 2000-04-07
22 PRIOR APPLICATION NUMBER: 31 80/159162
23 PRIOR FILING DATE: 2000-08-11
24 PRIOR APPLICATION NUMBER: 31 80/159162
25 NUMBER OF SEQ ID NOS: 7059
26 SOFTWARE: PatentIn ver. 3.0

```

Tue Apr 8 16:01:57 2003

us-09-647-749a-1.rapb

SEQ ID NO 6675
LENGTH: 332
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6675

Query Match 100.0% Score 25; DB 9; Length 332;
Best Local Similarity 100.0%; Pctd. No. 17e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
11111
DB 234 LESYT 238

RESULT 15
US-10-160-501-14
Sequence 14, Application US/20160501
Publication No. US20030059919A1

GENERAL INFORMATION:

APPLICANT: Meyers, Rachel E.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Kappeler-Liebermann, Rosana
TITLE OF INVENTION: NOVEL HUMAN 39228, 21956, 25856, 22244, 8701, 32263,
TITLE OF INVENTION: 50250, 55158, 47765, 62088, 50566, AND 48118
TITLE OF INVENTION: MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-250
CURRENT APPLICATION NUMBER: US/10/160,501

CURRENT FILING DATE: 2002-05-30

PRIOR APPLICATION NUMBER: US 09/838,573

PRIOR FILING DATE: 2002-04-18

PRIOR APPLICATION NUMBER: US 60/197,747

PRIOR FILING DATE: 2000-04-18

PRIOR APPLICATION NUMBER: US 09/870,133

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207,649

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/470,130

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207,640

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/862,535

PRIOR FILING DATE: 2001-05-21

PRIOR APPLICATION NUMBER: US 60/205,961

PRIOR FILING DATE: 2000-05-19

PRIOR APPLICATION NUMBER: US 09/870,383

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207,506

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/860,821

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: US 60/205,449

PRIOR FILING DATE: 2000-05-19

PRIOR APPLICATION NUMBER: US 09/870,110

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: US 60/207,650

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/907,509

PRIOR FILING DATE: 2001-07-16

PRIOR APPLICATION NUMBER: US 60/218,385

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 09/945,327

PRIOR FILING DATE: 2001-08-31

PRIOR APPLICATION NUMBER: US 60/229,425

PRIOR FILING DATE: 2000-08-31

PRIOR APPLICATION NUMBER: US 60/318,581

PRIOR FILING DATE: 2001-09-10

NUMBER OF SEQ ID NOS: 41
SOFTWARE: FastSeq Version 4.0
SEQ ID NO 14
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens

US-10-160-501-14

Query Match 100.0% Score 25; DB 9; Length 467;
Best Local Similarity 100.0%; Pctd. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESYT 5
11111
DB 248 LESYT 252

Search completed: April 8, 2004 12:06:20
Job time: 10.1667 secs

ORGANISM: Staphylococcus epidermidis
09-134-001C-3015

Query Match 80.6% Score 29; DB 4; Length 703;
est Local Similarity 83.3%; Pred. No. 2,1e+02;
atches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2 KEYS 7
1:1111

195 KDYFTS 200

ULT 4
08-856-841-9
Sequence 9, Application US/0881583
Patent No. 6218142

GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHER

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES

TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS P19

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY

STREET: 99 PARK AVENUE

CITY: NEW YORK

STATE: NY

COUNTRY: USA

ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC

COMPUTER: AT&T - IBM COMPATIBLE

OPERATING SYSTEM: MS-DOS Version 6.2

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856,841

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/214,659

FILING DATE: 16-MARCH-1994

APPLICATION NUMBER: US 07/917,096

FILING DATE: 4-AUGUST-1992

APPLICATION NUMBER: PCT/DE91/00106

FILING DATE: 8-FEBRUARY-1991

APPLICATION NUMBER: DE40038262

FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.

REGISTRATION NUMBER: 27,224

REFERENCE/DOCKET NUMBER: LYP-922-A

TELEPHONE: (212) 697-3355

TELEFAX: (212) 557-5635

TELEX: NONE

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 210

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE: PEPTIDE

DESCRIPTION: N/A

ANTI-SENSE: N/A

FRAGMENT TYPE: INTERNAL

ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE

IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE

POSITION IN GENOME: N/A

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD: amino acid analysis and

IDENTIFICATION METHOD: mass spectrometry

OTHER INFORMATION:

PUBLICATION INFORMATION:

AUTHORS: COSSART, Y. E.

AUTHORS: FIELD, A. M.

AUTHORS: GANT, B.

AUTHORS: WIDOWS, D.

TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA

JOURNAL: LANCET

VOLUME: 1

ISSUE: 72 - 73

PAGES: 1975

DATE: 1975

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 9:

PUBLICATION INFORMATION:

AUTHORS: MANIATIS, T.

AUTHORS: FRITSCH, E. F.

AUTHORS: SAMBROOK, J.

TITLE: MOLECULAR CLONING

JOURNAL: COLD SPRING HARBOR, NY

VOLUME:

ISSUE:

PAGES:

DATE: 1982

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 9:

PUBLICATION INFORMATION:

AUTHORS: SMITH, D. B.

AUTHORS: THOMPSON, K. S.

TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES

TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH

TITLE: GLUTATHIONE S- TRANSFERASE

JOURNAL: GENE

VOLUME:

ISSUE: 67

PAGES: 31 - 40

DATE: 1988

DOCUMENT NUMBER:

FILING DATE:

PUBLICATION DATE:

RELEVANT RESIDUES IN SEQ ID NO: 9:

US-08-856-841-9

Query Match 77.8% Score 27; DB 4; Length 210;

est Local Similarity 66.7%; Pred. No. 92;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 KEYS 6
1:1111

Db 194 KDYFT 199

RESULT 4

US-08-811-583-3

Sequence 3, Application US/0881583

Patent No. 6218142

GENERAL INFORMATION:

APPLICANT: Wasseneget, Michael

APPLICANT: Riedel, Leonhard

APPLICANT: Schiebel, Winfried

APPLICANT: Saenger, Heinz

TITLE OF INVENTION: NUCLEIC ACID MIMETICS ENCODING

TITLE OF INVENTION: POLYPEPTIDES HAVING THE ENZYMAIC ACTIVITY OF AN

TITLE OF INVENTION: RNA-DIRECTED RNA POLYMERASE (RDRP)

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: FISH & NEAVE

STREET: 1251 Avenue of the Americas

CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-647-749A
FILING DATE: 05-MAR-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Haley, James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MFS-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
IS-08-811-583-3

Query Match: 77.8% Score 28; DP 4; Length 218;
Best Local Similarity: 57.1%; Pred. No: 96;
Matches: 4; Conservative: 3; Mismatches: 0; Indels: 0; Gaps: 0

1 KEYS: 7
144 VEIPTN 150

RESULT 5
IS-08-856-841-15
Sequence: 15; Application US/08850841
Patent No: 6274407

GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OF POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-647-749A
FILING DATE: 05-MAR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08-856-841-15
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US-07-917-096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/JP91/00107
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9227-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TEXT: NONE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 227
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPYTHMA INFECTION)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSAFT, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 15:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: PRITSCH, F.F.
AUTHORS: SAMBROOK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 15:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S-TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 15:
IS-08-856-841-15

Query Match 77.8% Score 28 DB 4: Length 227;
 Best Local Similarity 66.7% Pred. NO. 1e+02;
 Matches 4: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

1 KEYFT 6
 111111
 193 VKDYFT 198

SULT 6
 -08-856-841-12
 Sequence 12, Application US/08856841
 Patent No. 6274307
 GENERAL INFORMATION:

APPLICANT: MANFRED MOTZ
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: CP POLYPEPTIDES FROM THE PARVAVIRUS P19
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROOKS BAIRD HAFNER & DELAUNTY
 STREET: 96 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS Version 6.2
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-08-957 841;
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US-08-214,658
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US 07/917,096
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PCT/DE91/00106
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE40038262
 FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.
 REGISTRATION NUMBER: 27,224
 REFERENCE/COCKET NUMBER: LKF-0222-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-3355
 TELEFAX: (212) 557-5635

TELEX: NONE
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:

LENGTH: 250
 TYPE: AMINO ACID
 TOPOLOGY: LINEAR
 MOLECULE TYPE:
 DESCRIPTION: PEPTIDE
 ANTI-SENSE: N/A
 HYPOTHETICAL: N/A
 FRAGMENT TYPE: INTERNAL
 ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
 ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
 IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
 POSITION IN GENOME: N/A
 FEATURE:
 NAME/KEY: N/A
 LOCATION: N/A
 IDENTIFICATION METHOD: amino acid analysis and
 IDENTIFICATION METHOD: mass spectrometry
 OTHER INFORMATION:
 PUBLICATION INFORMATION:

AUTHORS: COSSART, Y.E.
 AUTHORS: FIELD, A.M.
 AUTHORS: CANT, B.
 AUTHORS: WIDOWS, D.
 TITLE: PARVAVIRUS-LIKE PARTICLES IN HUMAN SERA
 JOURNAL: LANCET
 VOLUME: 1
 ISSUE:
 PAGES: 72 - 73
 DATE: 1975

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 12:

PUBLICATION INFORMATION:
 AUTHORS: MANIATIS, T.
 AUTHORS: FRITSCH, E.F.
 AUTHORS: SAMBROOK, J.
 TITLE: MOLECULAR CLONING
 JOURNAL: COLD SPRING HARBOR, NY
 VOLUME:
 ISSUE:
 PAGES:
 DATE: 1982

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 12:
 PUBLICATION INFORMATION:

AUTHORS: SMITH, D.B.
 AUTHORS: JOHNSON, K.S.
 TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
 TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
 TITLE: GLUTATHIONE S- TRANSFERASE
 JOURNAL: GENE
 VOLUME:
 ISSUE: 67
 PAGES: 31 - 40
 DATE: 1988

DOCUMENT NUMBER:
 FILING DATE:
 PUBLICATION DATE:
 RELEVANT RESIDUES IN SEQ ID NO: 12:
 US-08-856-841-12

Query Match 77.8% Score 28 DB 4: Length 250;
 Best Local Similarity 66.7% Pred. NO. 1e+02;
 Matches 4: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

1 KEYFT 6
 111111
 DB 200 VKDYFT 205

RESULT 7
 US-08-935-263-16

Sequence 16, Application US/08935263A
 Patent No. 6117669
 GENERAL INFORMATION:
 APPLICANT: Furuichi, Yasuhiro
 APPLICANT: Hoshino, Tatsuo
 APPLICANT: Kimura, Hitoshi
 APPLICANT: Kiyasu, Tatsuya
 APPLICANT: Nagashima, Yoshie
 TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
 FILE REFERENCE: Biotin Genes
 CURRENT APPLICATION NUMBER: US/08/935,263A
 CURRENT FILING DATE: 1997-09-22
 EARLIER APPLICATION NUMBER: EP 96115540.5
 EARLIER FILING DATE: 1996-09-27
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 16

LENGTH: 276
 TYPE: PRT
 ORGANISM: Kurthia sp.
 IS-08-935-263-16

Query Match 77.8% Score 28; DB 3; Length 276
 Best Local Similarity 57.1% Pred. No. 1,2e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0;

1 KEYFITS 7
 111111
 214 VEEFFTS 220

RESULT 9
 IS-09-594-185-16
 Sequence 16, Application US/08/054195
 Patent No. 6165388
 GENERAL INFORMATION:

APPLICANT: Furutachi, Yasuhiro
 APPLICANT: Hoshino, Tatsuo
 APPLICANT: Kimura, Hitoshi
 APPLICANT: Kiyasu, Tatsuya
 APPLICANT: Nagahashi, Yoshie
 TITLE OF INVENTION: PROTEIN BIOSYNTHETIC GENES
 FILE REFERENCE: Biotin Genes
 CURRENT APPLICATION NUMBER: US/09/094,185
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: 08/935,263
 PRIOR FILING DATE: 1997-09-22
 PRIOR APPLICATION NUMBER: EP 96115540.5
 PRIOR FILING DATE: 1996-09-27
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 16

LENGTH: 276
 TYPE: PRT
 ORGANISM: Kurthia sp.
 IS-09-594-185-16

Query Match 77.8% Score 28; DB 4; Length 276,
 Best Local Similarity 57.1% Pred. No. 1,2e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0;

1 KEYFITS 7
 111111
 214 VEEFFTS 220

RESULT 9
 S-08-605-284B-18
 Sequence 18, Application US/08/05284B
 Patent No. 6060271
 GENERAL INFORMATION:

APPLICANT: WALEWSKI, JOSE F
 APPLICANT: RECIO-PINTO, ESPERANZA
 TITLE OF INVENTION: VOLTAGE GATED SODIUM CHANNELS FROM
 TITLE OF INVENTION: HUMAN PERIPHERAL NERVE
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
 STREET: CLINTON SQUARE, P.O. BOX 1051
 CITY: ROCHESTER
 STATE: NEW YORK
 COUNTRY: USA

ZIP: 14603
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/052,84B

FILED DATE: 09-FEB-1996
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: BRAMAN, SUSAN J.
 REGISTRATION NUMBER: 34,103
 REFERENCE/EXCERPT NUMBER: 1996-09-01 (OFF D-1705)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 716-263-1636
 TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 311 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-605-284B-18

Query Match 77.8% Score 28; DB 4; Length 311;
 Best Local Similarity 71.4% Pred. No. 1,4e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 KEYFITS 7
 111111
 DB 147 KEYFITS 153

RESULT 10
 US-08-856-841-17
 Sequence 17, Application US/08/056841
 Patent No. 6274107
 GENERAL INFORMATION:

APPLICANT: ERWIN SOUTSCHER
 APPLICANT: MANFRED MOLT
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIOT HAFNER & DELAHUNTY
 STREET: 99 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS Version 5.0
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,841
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/014,694
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US 07/917,054
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PT/96/01,00104
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE40038262
 FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.
 REGISTRATION NUMBER: 27,224
 REFERENCE/EXCERPT NUMBER: LRP 92200A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 697-3355
 TELEFAX: (212) 557-5635
 TELEX: NONE

INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 387

TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOHEMICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPITHEMIA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
AUTHORS: SAMROCK, J.
AUTHORS: FRITSCH, E.F.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 17:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMROCK, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES:
DATE: 1982
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 15:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S. TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 17:
-08-856-841-17
Query Match 77.88, Score 28, DB 4, Length 387;
Best Local Similarity 66.78, Freq. No. 1.8e+02,
Matches 4, Conservative 2, Mismatches 0, Indels 0, Gaps 0;
1 KEYFT 6
1:1111
201 VKDYFT 206
SULT 11

US-08-856-841-19
Sequence 19, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: BROOKS HAITP HAFNER & DELAUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 05/08/214,654
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,094
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9227-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3455
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 486
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOHEMICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPITHEMIA INFECTIONOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDOWS, D.
AUTHORS: SAMROCK, J.
TITLE: MOLECULAR CLONING
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 19:
S-08-856-841-19

Query Match 77.88; Score 28; DB 4; Length 486;
Best Local Similarity 66.78; Pctd. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0.

1 KEKEYT 6
193 VKDYFT 198

RESULT 12
S-08-856-841-16
Sequence 16 Application US/08856841
Patent No. 6274307

GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVAVIRUS B13
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08-856-841
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08-8214-556
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US-07-917-036
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE0018262
FILING DATE: 8-FEBRUARY-1990

ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 29,224
REFERENCE: T-FILE NUMBER: 17-00000-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3155
TELEFAX: (212) 557-5635
TELEX: NONE

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:

LENGTH: 500
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (EPITHEMA INFECTIONUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:

NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:

PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WILKINS, D.
TITLE: PARVAVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 16:
PUBLICATION INFORMATION:
AUTHORS: MANIATIS, T.
AUTHORS: FRITSCH, E.F.
AUTHORS: SAMPOUR, J.
TITLE: MOLECULAR CLONING
JOURNAL: COLD SPRING HARBOR, NY
VOLUME:
ISSUE:
PAGES: 1982

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 16:
PUBLICATION INFORMATION:
AUTHORS: SMITH, D.B.
AUTHORS: JOHNSON, K.S.
TITLE: SINGLE STEP PURIFICATION OF POLYPEPTIDES
TITLE: EXPRESSED IN ESCHERICHIA COLI AS FUSIONS WITH
TITLE: GLUTATHIONE S-TRANSFERASE
JOURNAL: GENE
VOLUME:
ISSUE: 67
PAGES: 31 - 40
DATE: 1988

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 16:
US-08-856-841-16
Query Match 77.88; Score 28; DB 4; Length 500;
Best Local Similarity 66.78; Pctd. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0.

QY 1 KEKEYT 6
Db 201 VKDYFT 206

RESULT 13
S-08-856-841-18
Sequence 18 Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHER
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVAVIRUS B13
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE

CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/04/956,841
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/04/214,678
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/JP93/000106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LRR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3455
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ. ID NO.: 1A:
SEQUENCE CHARACTERISTICS:
LENGTH: 501
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A

FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIOSUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:

NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y. E.
AUTHORS: FLETH, A. M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: 1
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 18:
-08-856-841-18

Query Match 77.8% Score 28; DB 4; Length 501;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 IKEYFT 6
|||||

Db 202 VKDYFT 207

RESULT 14
US-09-413-452-2
Sequence 2, Application US/09413452
Patent No. 6083540
GENERAL INFORMATION:
APPLICANT: Christensen, T.
APPLICANT: Thorsoe, H.
APPLICANT: Kreiberg, J.
APPLICANT: Buchholt, H.
TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN
ACIDIC ENVIRONMENT WITH A HIGH-ESTER PECTIN
FILE REFERENCE: DYN014 001APC
CURRENT APPLICATION NUMBER: US/09/413,452
FILING DATE: 1999-10-06
EARLIER APPLICATION NUMBER: 08/983364
FILING DATE: 1998-05-18
EARLIER APPLICATION NUMBER: PCT/EP96/0451
FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 584
TYPE: PPT
ORGANISM: Citrus navelina var. class 1
US-09-413-452-2

Query Match 77.8% Score 28; DB 3; Length 584;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6
|||||
Db 4 IKEYFT 9

RESULT 15
US-09-413-068-2
Sequence 2, Application US/09413068
Patent No. 6268195
GENERAL INFORMATION:
APPLICANT: Christensen, T.
APPLICANT: Thorsoe, H.
APPLICANT: Kreiberg, J.
APPLICANT: Buchholt, H.
TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN
ACIDIC ENVIRONMENT WITH A HIGH-ESTER PECTIN
FILE REFERENCE: DYN014 001APC
CURRENT APPLICATION NUMBER: US/09/413,068
FILING DATE: 1999-06-06
EARLIER APPLICATION NUMBER: 08/983364
FILING DATE: 1998-05-18
EARLIER APPLICATION NUMBER: PCT/EP96/0451
FILING DATE: 1996-07-12
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 584
TYPE: PPT
ORGANISM: Citrus navelina var. class 1
US-09-413-068-2

Query Match 77.8% Score 28; DB 4; Length 584;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6
|||||
Db 4 IKEYFT 9

Tue Apr 8 16:02:02 2003

us-09-647-749a-2.ra1

Page 9

Search completed: April 8, 2003, 12:05:53
Job time : 23.5833 secs

|||||
367 IKEYFT 372

SUIT 2
CV-CHICK

PGCV-CHICK STANDARD: PRT: 3562 AA.

09053: 090945:

01-NOV-1997 (Rel. 35, Created)

01-NOV-1997 (Rel. 35, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

Version core protein precursor (Large fibroblast proteoglycan)

(Chondroitin sulfate proteoglycan core protein 2) (PG-M).

CP562.

Gallus gallus (Chicken).

Eukaryota; Metazoa; Chordata; Cladacea; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

Gallus.

NCBI_TaxID=9031;

(1)
SEQUENCE FROM N.A.

STRAIN-White Leghorn; TISSUE=Limb bud;

MEDLINE=93300846; PubMed=8314802;

Shinomura T., Nishida Y., Ito K., Kimura K.;

cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan

expressed during chondrogenesis in chick limb buds. Alternative

spliced multiforms of PG-M and their relationships to versican."

J. Biol. Chem. 268:14461-14469(1993).

-1- FUNCTION: May play a role in intercellular signaling and in

connecting cells with the extracellular matrix. May take part in

the regulation of cell motility, growth and differentiation. Binds

hyaluronic acid.

-1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.

-1- ALTERNATIVE PRODUCTS: At least 2 isoforms; V0 (shown here) and

V1; are produced by alternative splicing.

-1- TISSUE SPECIFICITY: Prechondrogenic condensation area of

developing limb buds.

-1- DEVELOPMENTAL STAGE: Disappears after the cartilage development

(by similarity).

-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

-1- SIMILARITY: CONTAINS 2 LINK DOMAINS.

-1- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.

-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

-1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.

This SWISS-Prot entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed, usage by and for commercial

entities requires a license agreement (See [http://www.isb.ac.uk/ncbi/](http://www.isb.ac.uk/ncbi/ncbi/)

or send an email to license@isb.slb.ch).

EMBL: X60226; CAA42787.1; .

EMBL: D13542; BA02742.1; .

HSSP: P00743; IELM.

InterPro: IPR000152; ASX-Hydroxyl.

InterPro: IPR000561; EGF-like.

InterPro: IPR000742; EGF-like.

InterPro: IPR001861; EGF-like.

InterPro: IPR003599; IG-1.

InterPro: IPR003606; IG-1.

InterPro: IPR001304; LECTIN_C.

InterPro: IPR000538; Link.

InterPro: IPR000436; Sushi_SCR_DOM.

pfam: PF000047; IGF_1.

pfam: PF00059; IGF_1.

pfam: PF00084; Sushi_1.

pfam: PF00193; Link_2.

ProDom: PD000918; Link_2.

SMART: SM00032; CCP_1.

SMART: SM00034; CLECT_1.

DR SMART: SM00179; EGF_CA_1.

DR SMART: SM00001; EGF-Like_1.

DR SMART: SM00409; IG_1.

DR SMART: SM00445; LINK_2.

DR PROSITE: PS00010; ASX-Hydroxyl_1.

DR PROSITE: PS00022; EGF-1_2.

DR PROSITE: PS01186; EGF-2_1.

DR PROSITE: PS01187; EGF-CA_1.

DR PROSITE: PS01241; LINK_2.

DR PROSITE: PS00615; C-TYPE-LECTIN_1.

DR PROSITE: PS00041; C-TYPE-LECTIN_2_1.

DR Glycoprotein: Proteoglycan; Lectin; Extracellular matrix; Sushi;

KM Signal: Repeat; EGF-like domain; Calcium; Immunoglobulin domain;

KW Hyaluronic acid; Alternative splicing.

KW SIGNAL

FT CHAIN 1 26

FT DOMAIN 27 3562

FT DOMAIN 37 136

FT DOMAIN 166 243

FT DOMAIN 264 345

FT DOMAIN 3254 3290

FT DOMAIN 3292 3328

FT DOMAIN 3341 3455

FT DOMAIN 3460 3518

FT DISULFID 44 129

FT DISULFID 171 242

FT DISULFID 195 216

FT DISULFID 269 344

FT DISULFID 293 314

FT DISULFID 3258 3269

FT DISULFID 3263 3278

FT DISULFID 3280 3289

FT DISULFID 3301 3307

FT DISULFID 3316 3316

FT DISULFID 3318 3327

FT DISULFID 3334 3345

FT DISULFID 3362 3454

FT DISULFID 3430 3446

FT DISULFID 3461 3504

FT DISULFID 3490 3517

FT CARBOHYD 152 162

FT CARBOHYD 235 235

FT CARBOHYD 329 329

FT CARBOHYD 529 529

FT CARBOHYD 709 709

FT CARBOHYD 948 948

FT CARBOHYD 1406 1406

FT CARBOHYD 1479 1479

FT CARBOHYD 1523 1523

FT CARBOHYD 1530 1530

FT CARBOHYD 1545 1545

FT CARBOHYD 1751 1751

FT CARBOHYD 1968 1968

FT CARBOHYD 2088 2088

FT CARBOHYD 2089 2089

FT CARBOHYD 2507 2507

FT CARBOHYD 2642 2642

FT CARBOHYD 2679 2679

FT CARBOHYD 2748 2748

FT CARBOHYD 2762 2762

FT CARBOHYD 3069 3069

FT CARBOHYD 3194 3194

FT CARBOHYD 3232 3232

FT CARBOHYD 3545 3545

FT VARSPPLIC 485 1411

SV SEQUENCE 3562 AA; 388078 MW; 945565E8516222 CP064;

Query Match

Best Local Similarity

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 KEYS 7

DB 1355 KEYS 1360

Score 32; DB 1; Length 3562;

Pred. No. 1.7e+02;

88.9%;

```

RESULT 3
TC1_DICD1
D RTCL_DICD1 STANDARD: PRT: 433 AA.
C 015746:
T 15-DEC-1998 (Rel. 37, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
E Probable RNA 3'-terminal phosphate cyclase (EC 6.5.1.4) (PNA-3'-phosphate cyclase) (RNA cyclase).
N R2037.
S Dictyostelium discoideum (Slime mold)
C Eukaryota: Mycetozoa: Elenchozoa: Dictyostelium
X NCBI_TaxID=44689.
P SEQUENCE FROM N.A.
(1)
N STRAIN-AX3:
A Loomis W.F., Tranter N.
C Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases
C -1- FUNCTION: CATALYZES THE CONVERSION OF 3'-PHOSPHATE TO A 3',3'-CYCLIC PHOSPHODIESTER AT THE END OF RNA. THE MECHANISM OF ACTION OF THE ENZYME OCCURS IN 3 STEPS: (A) ADENYLATION OF THE ENZYME BY ATP; (B) THE ENZYME ACTS ON RNA-N3-P TO PRODUCE RNA-N3/PP5'A; (C) A NON CATALYTIC NUCLEOPHILIC ATTACK BY THE ADJACENT 2' HYDROXYL ON THE PHOSPHORUS IN THE FIRST STEP LINKAGE TO PRODUCE THE CYCLIC END-PRODUCT. THE BIOLOGICAL ROLE OF THIS ENZYME IS UNKNOWN BUT IT IS LIKELY TO FUNCTION IN SOME ASPECTS OF CELLULAR RNA PROCESSING (BY SIMILARITY)
C -1- CATALYTIC ACTIVITY: ATP + RNA 3'-terminal-phosphate -> AMP + diphosphate + RNA terminal-2',3'-cyclic-phosphate.
C -1- SUBCELLULAR LOCATION: Nuclear; nucleoplasm (By similarity).
C -1- SIMILARITY: BELONGS TO THE RNA 3'-TERMINAL CYCLASE FAMILY. SUBFAMILY 1.
C -1- CAUTION: THE C-TERMINAL SECTION OF THIS PROTEIN COULD BE INCOMPLETE AND THE REAL PROTEIN IS PROBABLY SHORTER.
C THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch) or send an email to license@sib-sib.ch.
C EMBL: AF020279; AAB70847.1; -.
C Dictyob: D00222; -.
C InterPro: IPR000228; RNA3'_term_cycl.
C Pfam: PF01137; RNC1.
C ProSite: PS01289; RNC1.
C Ligase: Nuclear protein.
W ACT_SITE 354 354
Q SEQUENCE 433 AA: 47110 MW: 53496044458069 GRC64:

Query Match 86.18: Score 31; DB 1; Length 433;
Best Local Similarity 85.78; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KEYSPTS 7
111111
b 229 LKEYPTS 234

```

```

OC Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=9526475; PubMed=7744966;
RA Gersch L.C., Corkendorff T.C., Collier N.
RT "A conditional allele of the novel repeat-containing yeast nucleoporin NUP159 causes both rapid cessation of mRNA export and reversible clustering of nuclear pore complexes."
RL J. Cell Biol. 129:939-956(1995).
RN (2)
RP SEQUENCE FROM N.A.
P STRAIN=SE887; AF072;
RA Baird B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Chutcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagals K., Jones M.,
RA Louis E., Lyne G., Moule S., Moule T., Odell C., Pearson D.,
RA Patandream M.A., Piles L., Powley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.
C Submitted (Sep-1994) to the EMBL/GenBank/DBJ databases.
C -1- FUNCTION: MAY PLAY A DIRECT ROLE IN NUCLEOTIDYLTRANSFERASE EXPORT OF RNA AND ALSO IN PROTEIN IMPORT.
C -1- SUBUNIT: INTERACTS WITH DBP5.
C -1- SUBCELLULAR LOCATION: Nuclear pore complex.
C -1- DOMAIN: CONTAINS A CENTRAL REGION WITH REPEATS THAT ARE X-X-F-G. PRIMARILY S-A-F-G AND P-S-F-G. SOME OF THESE REPEATS ARE EMBEDED WITH IN A NEARLY PERFECT 26 AMINO ACID TANDEM REPEATS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC EMBL: L40634; AAC41652.1; -.
CC TR EMBL: Z8125; CAAB6265.1; -.
CC DR PIR: S48457; S48457.
CC S300: S0003377; NUP159.
CC InterPro: IPR004326; Nucleoporin_FG.
CC Pfam: PF01494; Nucleoporin_FG.
CC K1: K1.
CC KW Nuclear protein; Transp.; Repeat.
CC FT DOMAIN 518 617
CC FT REPEAT 518 543 1.
CC FT REPEAT 544 569 2.
CC FT REPEAT 570 595 3.
CC FT REPEAT 596 617 4.
CC SEQUENCE 1460 AA: 158907 MW: 444945237858205 GRC64:

Query Match 86.18: Score 41; DB 1; Length 1460;
Best Local Similarity 71.48; Pred. No. 11e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 KEYSPTS 7
111111
Db 1190 LKEYPTS 1196

```

```

RESULT 5
PTKA_ECOLI
ID PTKA_ECOLI STANDARD: PRT: 150 AA.
AC P37187; P76413;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PTS system, galactitol-specific IIA component (PtkA-GAT) (galactitol-
permease IIA component) (Phosphotransferase enzyme II A component)
DE (EC 2.7.1.69).
DE GATA OR B2094 OR Z3257 OR FCS2897.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OS

```


Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae,

Escherichia

NCBI_TaxID=562, 83334;

(1)

SEQUENCE FROM N.A.

STRAIN=EC3132;

MEDLINE=95290477; PubMed=7775623;

(2)

SEQUENCE OF THE gat operon for galactitol utilization from a

wild-type strain EC3132 of *Escherichia coli*;

Biochim. Biophys. Acta 1262:69-72(1995).

(3)

SEQUENCE FROM N.A.

STRAIN=K12 / MG1655;

MEDLINE=97426617; PubMed=9278503;

Blaetter F.R., Plunkett G. III, Bloch C.A., Perma N.T., Burland V.,

Riley M., Collado-Vides J., Glasner J.D., Mayhew G.F.,

Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

Mau B., Shao Y.;

"The complete genome sequence of *Escherichia coli* K-12."

Science 277:1232-1238(1997).

(4)

SEQUENCE FROM N.A.

STRAIN=O157:H7 / RMD 0509952;

MEDLINE=21156231; PubMed=11258796;

Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

Han C.-G., Ohtsuo E., Nakayama K., Murata T., Tanaka M., Tohe T.,

Iida T., Takami H., Honda T., Sasakawa G., Ogasawara N., Yasunaga T.,

Kudoh S., Shiba T., Hattori M., Shindagawa H.;

"Complete genome sequence of enterohemorrhagic *Escherichia coli*

O157:H7 and genomic comparison with a laboratory strain K-12."

DNA Res. 8:11-22(2001).

(5)

FUNCTION: THIS IS A COMPONENT OF THE PHOSPHONOPYRUVATE-DEPENDENT

SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE

TRANSPORT SYSTEM. THE TCB DOMAIN CONTAINS THE SUGAR BINDING SITE

AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY

PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HFP); THE TRANSFERS ITS

PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO

THE SUGAR.

(6)

CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein

histidine + sugar phosphate

(7)

SUBCELLULAR LOCATION: Cytoplasmic

Phosphorylation: Galactitol metabolism; (complete) proteome.

KT CONFLICT 124 124 D -> E (IN REF. 1).

ST SEQUENCE 150 AA: 16907 MW: 1A760076B5FB61D C6C64;

Query Match

Best Local Similarity 83.3% Score 100 DB 1; Length 150;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYEF 6

DB 137 IKEYEF 142

RESULT 6

PYBL_PEA STANDARD: PRT: 405 AA.

ID PYBL_PEA

AC 043086;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Aspartate carboxyltransferase 1, chloroplast precursor (EC 2.1.1.2)

DE (Aspartate transcarbamylase 1) (AtCase 1).

GN PYBL1

OS Pisum sativum (Garden pea).

OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

OC Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, Rosidae;

OC eurosids I, Fabales, Fabaceae, Papilionoideae, Viciae; Pisum.

OX NCBI_TaxID=3888;

RN (1)

RP SEQUENCE FROM N.A.

STRAIN=CV; Mando; TISSUE=Leaf;

EX MEDLINE=94402175; PubMed=8029359;

RA Williamson G.L., Slocum R.D.;

RT "Molecular cloning and characterization of the pybl1 and pybl2 genes

encoding aspartate transcarbamylase in pea (*Pisum sativum* L.)."

RL Plant Physiol. 105:377-384(1994).

CC -1- CATALYTIC ACTIVITY: Carboxyl phosphate + L-aspartate = phosphate

CC -1- ENZYME REGULATION: Allosterically regulated by UMP.

CC -1- PATHWAY: Pyrimidine biosynthesis; second step.

CC -1- SUBUNIT: HOMOTRIMER (POTENTIAL)

CC -1- SIMILARITY: BELONGS TO THE ATC4-8/ATC4S5 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch>).

CC or send an email to license@sib-sib.ch.

CC EMBL: M96981; AAA62443.1;

DR HSSP: P00479; 3CSU.

DR InterPro: IPR002039; Asp/Om/Prot.

DR InterPro: IPR002082; Asp_carbamyltransf.

DR Pfam: PF00185; OTCace.1.

DR Pfam: PF02729; OTCace.N.1.

DR PRINTS: P00100; AOTASF.

DR TRIGRAMS: TRIGR0670; asp_carb.tr.1.

DR PROSITE: PS00097; CARBAMOYLTRANSFERASE.1.

KW Pyrimidine biosynthesis; Transferrin; Chloroplast; Transit peptide;

KW Multigene family.

FT TRANSIT

FT CHAIN

FT SEQUENCE

386 AA: 42617 MW: A9440F45474E29F4 CR664;

Query Match

Best Local Similarity 83.3% Score 100 DB 1; Length 386;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IKEYFIS 7

DB 275 IKEYFIS 281

RESULT 7
 ANCL_PIRSP STANDARD: PRI: 569 AA.
 D MANC_PIRSP
 C P55298;
 T 01-OCT-1996 (Rel. 34, Created)
 T 01-OCT-1996 (Rel. 34, Last sequence update)
 T 15-JUL-1998 (Rel. 36, Last annotation update)
 E Mannan endo-1,4-beta-mannosidase C precursor (EC 3.2.1.78) (Beta-mannanase C) (1,4-beta-D-mannan mannanhydrolase C).
 MANC.
 S Pteromyces sp.
 C Eukaryota, Fungi: Chytridiomycota: Neocallimastixales;
 C Neocallimastixaceae: Pteromyces.
 X NCBI_TaxID=45796;
 N [1]
 P SEQUENCE FROM N.A.
 MEDLINE=96313314; PubMed=8768520;
 A Millward-Sadler S.J., Hall J., Black G.W., Hazlewood G.P., Gilbert H.J.;
 T "Evidence that the Pteromyces gene family encoding endo-1,4-mannanases arose through gene duplication."
 L FEBS Microbiol. Lett. 141:183-188(1996).
 C -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.
 C -1- SIMILARITY: THE 39 AA REPEATS ARE SIMILAR TO THOSE IN MANA, MANC AND XYNA; AND TO THOSE OF N.PATRICIAPOM XYNA
 C -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.
 C THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcements/> or send an email to license@isb-sib.ch).
 C -----
 R EMBL: X97520; CAA66134.1;
 R InterPro: IPR002883; CBD_5;
 R InterPro: IPR008085; Glyco_hydro_26;
 R InterPro: IPR001230; PreyL_site;
 R Pfam: PF02013; CBM_10; 2;
 R Pfam: PF02156; Glyco_hydro_26; 1;
 R PRINTS: PR00739; GLHYDRLASE26;
 W HydroLase: Glycosidase; Signal: Multigene family; Repeat.
 T SIGNAL 1 18 POTENTIAL.
 T CHAIN 19 569 MANNA ENDO-1,4-BETA-MANNOSIDASE C.
 T DOMAIN 490 569 2 X 39 AA APPROXIMATE REPEATS.
 T REPEAT 488 526 1;
 T REPEAT 531 569 2;
 T DOMAIN 473 477 POLY-ASN.
 T DOMAIN 480 486 POLY-ASN.
 Q SEQUENCE 569 AA: 64115 MW: 19277764E18328B5 CRC64;
 Query Match 83.3% Score 30; DB 1; Length 569;
 Best Local Similarity R 74; Pred No. 66;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Y 1 KEYPT 6
 :|||||
 b 444 KEYPT 449
 RESULT 8
 ANB_PIRSP STANDARD: PRI: 571 AA.
 D MANB_PIRSP
 C P55297;
 T 01-OCT-1996 (Rel. 34, Created)
 T 01-OCT-1996 (Rel. 34, Last sequence update)
 T 15-JUL-1998 (Rel. 36, Last annotation update)
 E Mannan endo-1,4-beta-mannosidase B precursor (EC 3.2.1.78) (Beta-

mannase B) (1,4-beta-D-mannan mannanhydrolase B).
 MANB.
 S Pteromyces sp.
 C Eukaryota, Fungi: Chytridiomycota: Neocallimastixales;
 C Neocallimastixaceae: Pteromyces.
 X NCBI_TaxID=45796;
 N [1]
 P SEQUENCE FROM N.A.
 MEDLINE=96313314; PubMed=8768520;
 A Millward-Sadler S.J., Hall J., Black G.W., Hazlewood G.P., Gilbert H.J.;
 T "Evidence that the Pteromyces gene family encoding endo-1,4-mannanases arose through gene duplication."
 L FEBS Microbiol. Lett. 141:183-188(1996).
 C -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-beta-D-mannosidic linkages in mannans, galactomannans, glucomannans, and galactoglucomannans.
 C -1- SIMILARITY: THE 39 AA REPEATS ARE SIMILAR TO THOSE IN MANA, MANC AND XYNA; AND TO THOSE OF N.PATRICIAPOM XYNA.
 C -1- SIMILARITY: BELONGS TO FAMILY 26 OF GLYCOSYL HYDROLASES.
 C THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcements/> or send an email to license@isb-sib.ch).
 C -----
 R EMBL: X97408; CAA66061.1;
 R InterPro: IPR002883; CBD_5;
 R InterPro: IPR008085; Glyco_hydro_26;
 R InterPro: IPR001230; PreyL_site;
 R Pfam: PF02013; CBM_10; 2;
 R Pfam: PF02156; Glyco_hydro_26; 1;
 R PRINTS: PR00739; GLHYDRLASE26;
 W HydroLase: Glycosidase; Signal: Multigene family; Repeat.
 T SIGNAL 1 19 POTENTIAL.
 T CHAIN 20 571 MANNA ENDO-1,4-BETA-MANNOSIDASE B.
 T DOMAIN 490 571 2 X 39 AA APPROXIMATE REPEATS.
 T REPEAT 490 527 1;
 T REPEAT 533 571 2;
 T DOMAIN 475 479 POLY-ASN.
 T DOMAIN 482 488 POLY-ASN.
 Q SEQUENCE 571 AA: 64397 MW: B13944581FAA90AA CRC64;
 Query Match 83.3% Score 40; DB 1; Length 571;
 Best Local Similarity R 74; Pred No. 66;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Y 1 KEYPT 6
 :|||||
 Db 446 KEYPT 451
 RESULT 9
 YKA9_YEAST STANDARD: PRI: 248 AA.
 ID YKA9_YEAST
 AC P33201;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 27.1 kDa protein from YAP1 intergenic region.
 GN YK1009W OR YK1160.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota, Fungi: Ascomycota: Saccharomycotina: Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 N [1]
 P SEQUENCE FROM N.A.
 MEDLINE=93127732; PubMed=1481574;
 RX Pascolo S., Glazvini M., Beyer J., Collaenx L., Thierry A., Dujon B.;
 FA "The sequence of a 9.3 kb segment located on the left arm of the

yeast chromosome XI reveals five open reading frames including the CCI1 gene and putative products related to MTO2 and to the ribosomal protein L10.¹

[2]

SEQUENCE OF 190-236 FROM N.A.
MEDLINE-93255900; PubMed-9466747;
Boyer J., Pascolo S., Richard G.F., Dujon B.:
"Sequence of a 7.8 kb segment on the left arm of yeast chromosome XI reveals four open reading frames, including the CCI1 gene, an intron-containing gene and a gene encoding a homology to the mammalian UOG-1 gene."
Yeast 9:279-287(1993).

-1- SIMILARITY: TO YEAST RIBOSOMAL PROTEIN L10.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch>) or send an email to license@isb-sib.ch).

EMBL: S53418; AAB24904.1; -
EMBL: X61398; -; NOT_ANNOTATED_CDS.
EMBL: Z28009; CAAB1844.1; -
EMBL: S59773; AAC60552.1; -
PIR: S30013; S30013.
SCD: S0001492; YKL009W.
InterPro: IPR001790; Ribosomal_L10.
Pfam: PF00466; Ribosomal_L10. 1.
Hypothetical protein.

SEQUENCE 236 AA. 27058 MW. 39CFEEFEC09F427 CRCK4;

Query Match 80.6%; Score 29; DB 1; Length 236;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 IKEYPTS 7
:|||||
112 VKETFS 118

SOLT 10
CA_THEME STANDARD; PRT: 338 AA.
HRCVSVS
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Heat-inducible transcription repressor hrcA homolog
HRCV OR TM0851.
Thermotoga maritima
Bacteria; Thermotogae; Thermotogae (class); Thermotogales.
Thermotogaceae; Thermotoga.
NCBI_TaxID=2336;
[1]

SEQUENCE FROM N.A.
STRAIN-MSB8 / DSM 3109;
MEDLINE-99287316; PubMed-10360571;
Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
Halt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.P., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salberg S.L., Smith H.O., Venter J.C., Fraser C.M.:
"Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima."
Nature 399:323-329(1999).

-1- FUNCTION: NEGATIVE REGULATOR OF CLASS I HEAT SHOCK GENES (GPEE-DNAJ AND GPELS OPERONS). PREVENTS HEAT-SHOCK INDUCTION OF THESE OPERONS (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE HRCV FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch>)
CC or send an email to license@isb-sib.ch).

EMBL: AE001751; AAD35933.1; -
DR TIGR: TM0851;
DR InterPro: IPR002571; HrcA.
PFam: PF01628; HrcA; 1.
DE TIGRFAWS: TIGR00331; hrcA; 1.
KW Transcription regulation; Repressor; Heat shock; Complete proteome.
SQ SEQUENCE 338 AA. 39306 MW. 02A05A546731AE CRCK4;

Query Match 80.6%; Score 79; DB 1; Length 338;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 IKEYPTS 7
:|||||
328 LSEYFS 334

RESULT 11
ACBC_ACTS5 STANDARD; PRT: 342 AA.
ID ACBC_ACTS5
AC Q9ZAE3; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
RT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-epi-5-epi-valiolone synthase (EC 4.2.3.-) (Sedo-heptulose 7-
DE phosphate cyclase).
GN ACBC.
OS Actinoplanes sp. (strain 50/110).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteriales;
OC Actinomycetales; Micromonosporineae; Micromonosporaceae;
OC Actinoplanes.
OX NCBI_TaxID=1871;
RN [1]

SEQUENCE FROM N.A.
MEDLINE-99214139; PubMed-10196166;
RX Stratmann A., Mahmud T., Lee S., Distler J., Floss H.G.,
RA Piepersberg W.:
"The AcBC protein from Actinoplanes species is a C7-cyclitol synthase
RT related to 3-dehydroquinate synthases and is involved in the
RT biosynthesis of the alpha-glucosidase inhibitor arabinose."
RL J. Biol. Chem. 274:10889-10896(1999).

CC -1- FUNCTION: C7-cyclitol synthase using sedo-heptulose 7-phosphate,
CC but not ido-heptulose 7-phosphate and 3-deoxy-arabino-
CC heptulose 7-phosphate, as a substrate. Involved in the
CC biosynthesis of the acarbose moiety of the alpha-glucosidase
CC inhibitor acarbose.
CC -1- CATALYTIC ACTIVITY: Sedo-heptulose 7-phosphate + 2-epi-5-epi-
CC -1- VALIOLONE + phosphate.
CC -1- COFACTOR: NAD and a divalent metal cation such as cobalt.
CC -1- SIMILARITY: BELONGS TO THE DEHYDROQUINATE SYNTHASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch>)
CC or send an email to license@isb-sib.ch).

EMBL: Y18523; CAA77208.1; -
DR InterPro: IPR002458; DHQ_synthase.
DR Pfam: PF01761; DHQ_synthase; 1.
KW Lyase; NAD; Cobalt.
SQ SEQUENCE 382 AA. 42181 MW. EDA70004A149A84 CRCK4;

SOLT 14
 RC_CLOPE STANDARD: PRI: 620 AA.
 08XN11:
 15-JUN-2002 (Rel. 41, Created)
 15-JUN-2002 (Rel. 41, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 Exonuclease ABC subunit C.
 UVRC OR CPE0352.
 Clostridium perfringens.
 Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium
 NCBI_TaxID=1502;
 [1]
 SEQUENCE FROM N.A.
 STRAIN-13 / Type A:
 PubMed=11792842;
 Shimizu T., Ohnishi K., Hirakawa H., Ohshima K., Yamashita A.,
 Shiba T., Ogasawara N., Hattori M., Kohara S., Hayashi H.;
 *Complete genome sequence of Clostridium perfringens, an anaerobic
 flesh-eater.
 Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE
 UVRA-UVRC COMPLEX, DISPLACING UVRA, AND THE DAMAGED DNA STRAND IS
 NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
 -1- SUBUNIT: Consists of three subunits: uvra, uvrb and uvrc.
 -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 -1- SIMILARITY: BELONGS TO THE UVRC FAMILY.
 -1- SIMILARITY: CONTAINS 1 TYP DOMAIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.slb.ch/announce/>
 or send an email to license@isb.slb.ch).

 EMBL: AP003186; BAB80058.1;
 InterPro: IPR000445; HhH.
 InterPro: IPR001943; UVRC.
 InterPro: IPR004791; UVRC.
 InterPro: IPR001162; UVRC.
 InterPro: IPR000305; UVRC_N.
 Pfam: PF01541; Exclendo_N.1.
 Pfam: PF02151; UVRC.1.
 ProDom: PD005870; UVRC_C.1.
 SMART: SM00465; GYRC.1.
 TIGRFAMs: TIGR00194; uvrc.1.
 PROSITE: PS0151; UVRC.1
 SOS response; Excision nuclease; DNA repair; Complete proteome.
 DOMAIN 204 239
 SEQUENCE 620 AA: 71390 MW: 66144DE9C3849219 QKCG4;
 Query Match 80.6%; Score 29; DB 1; Length 620;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 KEYFITS 7
 1111111
 614 IKOYSS 620
 SOLT 15
 81_YEAST STANDARD: PRT: 1562 AA.
 004781: 004029;
 01-NOV-1987 (Rel. 35, Created)
 01-NOV-1987 (Rel. 35, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Hypothetical 180.2 kDa protein in FAA4-HOR7 intergenic region.

GN YMP247C OR YMP408 ORC OR YMP920.01C
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 [1]
 PP SEQUENCE OF 1-956 FROM N.A.
 RC STRAIN-S288c / AB972;
 RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 RN SEQUENCE OF 950-1562 FROM N.A.
 PC STRAIN-S288c / AB972;
 RA Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -1- SIMILARITY: STRONG, TO HUMAN ZNF294.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.slb.ch/announce/>
 or send an email to license@isb.slb.ch).
 CC -----
 DR EMBL: Z48639; CAAB8574.1;
 DR EMBL: Z48756; CAAB8657.1;
 TR S50: S0004861; YMP247C.
 DR InterPro: IPR001841; ZnF_ring.
 DR SMART: SM00184; RING.1.
 DR PROSITE: PS00518; ZF_RING_1; FALSE_NRG.
 DR PROSITE: PS00089; ZF_RING_2; 1.
 KW Hypothetical protein; Zinc-finger.
 FT ZN-FING 1508 1555
 SQ SEQUENCE 1562 AA: 180185 MW: 97AC55E281362305 CRCG4;
 Query Match 80.6%; Score 29; DB 1; Length 1562;
 Best Local Similarity 71.4%; Pred. No. 3.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 1 KEYFITS 7
 1111111
 Db 462 LEDYFYS 468
 Search completed: April 8, 2003, 12:02:17
 Job time : 19.9167 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

on 01: April 8, 2003, 11:59:15 (search time 30.9167 seconds)

(without alignments)
21.766 Million cell updates/sec

sequence: 1 IKEYFTS 7

scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

searched: 283224 seqs, 96134422 residues

total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

2: IKEYFTS
3: IKEYFTS
4: IKEYFTS

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------------|
| 1 | 32 | 88.9 | 278 | E72425 | dihydropyruvate synthase |
| 2 | 32 | 88.9 | 406 | T32922 | hypothetical prote |
| 3 | 32 | 88.9 | 416 | S48411 | CMP-N-acetylneuram |
| 4 | 32 | 88.9 | 469 | S55167 | WEE2-dependent sig |
| 5 | 32 | 88.9 | 577 | A57469 | CMP-N-acetylneuram |
| 6 | 32 | 88.9 | 620 | E90575 | lipoprotein lipase |
| 7 | 32 | 88.9 | 3562 | A47171 | chondroitin sulfate |
| 8 | 31 | 86.1 | 260 | T50109 | gastrulation prote |
| 9 | 31 | 86.1 | 263 | T51235 | small protein - 2 |
| 10 | 31 | 86.1 | 1460 | S48457 | nucleoporin RNT7 - |
| 11 | 30 | 83.3 | 150 | S55903 | phosphotransferase |
| 12 | 30 | 83.3 | 150 | E64376 | pts system, galact |
| 13 | 30 | 83.3 | 150 | AV0991 | pts system, galact |
| 14 | 30 | 83.3 | 150 | C85836 | pts system, galact |
| 15 | 30 | 83.3 | 233 | S47625 | aspartate carbamoy |
| 16 | 30 | 83.3 | 386 | T06484 | aspartate carbamoy |
| 17 | 30 | 83.3 | 402 | H86649 | thiamosyltransfera |
| 18 | 30 | 83.3 | 487 | T27363 | hypothetical prote |
| 19 | 30 | 83.3 | 524 | T44908 | hypothetical prote |
| 20 | 30 | 83.3 | 524 | C07703 | hypothetical prote |
| 21 | 30 | 83.3 | 573 | T05338 | hypothetical prote |
| 22 | 30 | 83.3 | 598 | T93824 | probable terpene s |
| 23 | 30 | 83.3 | 1090 | H86806 | hypothetical prote |
| 24 | 29 | 80.6 | 92 | A62474 | hypothetical prote |
| 25 | 29 | 80.6 | 136 | C71661 | hypothetical prote |
| 26 | 29 | 80.6 | 176 | F91850 | ferric uptake regu |
| 27 | 29 | 80.6 | 199 | E89972 | hypothetical prote |
| 28 | 29 | 80.6 | 214 | H97767 | thiol, disulfide in |
| 29 | 29 | 80.6 | 240 | T01601 | probable secreted |

| | | | | | |
|----|----|------|------|--------|--------------------|
| 30 | 29 | 80.6 | 236 | E29013 | hypothetical prote |
| 31 | 29 | 80.6 | 255 | C95997 | probable glucose-1 |
| 32 | 29 | 80.6 | 293 | FA6257 | hypothetical prote |
| 33 | 29 | 80.6 | 303 | B69903 | aromatic metabolit |
| 34 | 29 | 80.6 | 338 | U73327 | heat shock operon |
| 35 | 29 | 80.6 | 360 | T42321 | complement control |
| 36 | 29 | 80.6 | 402 | S51791 | Drosophila translo |
| 37 | 29 | 80.6 | 406 | F70019 | hlf protein homol |
| 38 | 29 | 80.6 | 435 | T30114 | hypothetical prote |
| 39 | 29 | 80.6 | 524 | D82944 | hypothetical membr |
| 40 | 29 | 80.6 | 701 | F90038 | hypothetical prote |
| 41 | 29 | 80.6 | 704 | H82381 | toxin secretion AT |
| 42 | 29 | 80.6 | 764 | E95252 | conserved hypothet |
| 43 | 29 | 80.6 | 764 | C98117 | probable DNA-dirac |
| 44 | 29 | 80.6 | 1155 | T49421 | CCH2 type zinc fin |
| 45 | 29 | 80.6 | 1236 | T40120 | |

ALIGNMENTS

RESULT 1

E72425
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence-revision 11-Jun-1999 #text-change 02-Sep-2000
C:Accession: E72425
R:Neilsen, K.E.; Clayton, R.A.; Gill, S.P.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 359, 323-329, 1999
A>Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72909 M0109424700 TMD-1042471
A:Accession: E72425
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-278 <ARN>
A:Cross-references: GB:AE001691; GB:AE005512; NID:94960517; FIDN:AND35134.1; PID:9498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0040
C:Superfamily: dihydropyruvate synthase, dihydropyruvate synthase homology
C:Keywords: folate biosynthesis; transforase
F:18-358/Domain: dihydropyruvate synthase homology -DBS>

Query Match 88.9% Score 32 DB 2 Length 278;
Best Local Similarity 100.0% P-val No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYFT 6
|||||

Db 163 IKEYFT 168

RESULT 2

T32922
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence-revision 29-Oct-1999 #text-change 20-Oct-1999
C:Accession: T32922
R:Madson, C.; Graves, T.; Blair, T.
Submitted to the EMBL Data Library, January 1998
A:Reference number: 231247
A:Accession: T32922
A>Status: preliminary; translated from cDNA/EMBL/GenBank
A:Molecule type: DNA
A:Residues: 1-406 <MAT>
A:Cross-references: EMBL:AF047000; FIDN:ZAR7572.1; GSTRB:CH00019; CEST:K09H9.5
A:Experimental source: strain Bristol N. clone K09H9
C:Genetics:
A:Gene: CESP-K09H9.5
A:Map position: 1

Best Local Similarity 85.7% Pred. No. 66;
Matches 6: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Y 1 KEYFITS 7
|||||
b 250 KEYFITS 255

RESULT 7

hondrolin sulfate proteoglycan pg-M core protein - chicken

C:Species: Gallus gallus (chicken)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #extl_change 21-Jul-2000

C:Accession: A47171

C:Shimomura, T.; Nishida, Y.; Ito, K.; Kimata, K.

C:Biol. Chem. 268, 14461-14469, 1993

C:Title: cDNA cloning of pg-M, a large chondroitin sulfate proteoglycan expressed during

C:Reference number: A47171; MIMD:93300846; PMID:8314802

C:Accession: A47171

C:Status: preliminary

C:Molecule type: nucleic acid

C:Residues: 1-3562 <SH>

C:Cross-references: GB:D13542; MID:9331643; PIDN:BA02742.1; PID:9391644

C:Experimental source: Stage 22-23 developing limb buds

C:Note: sequence extracted from NCBI backbone (NCBI:134456, NCBI:134457)

C:Superfamily: chicken chondroitin sulfate proteoglycan pg-M core protein C-type lectin

C:166-235/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:124-147/Domain: Link protein repeat homology (LPR)

C:Species: Brachydanio rerio (zebra fish)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #extl_change 31-Jan-1997

C:Accession: J51225

C:Thisse, C.; Thisse, B.; Schilling, T.F.; Postlethwait, J.H.

C:Development 119, 1203-1215, 1993

C:Title: Structure of the zebrafish sna-1 gene and its expression in wild-type, spad

C:Reference number: J51225; MIMD:941954; PMID:840683

C:Accession: J51225

C:Status: preliminary; translated from GB/EMBL/DBJ

C:Molecule type: mRNA

C:Residues: 1-263 <TH>

C:Cross-references: GB:568799; MID:954544; PID:9545450

C:Genetics: sna-1

C:Gene: sna-1

C:Query Match

Best Local Similarity 86.1% Score 31; DB 2; Length 263;
Matches 5: Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYFITS 7
|||||

b 7 VKEYFITS 13

RESULT 10

hondrolin RAT7 yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 12-Nov-1994 #sequence_revision 12-Mar-1999

C:Accession: S48457

C:Reference number: S48457

C:Accession: S48457

C:Molecule type: DNA

C:Residues: 1-1460 <BCW>

C:Cross-references: GB:247437; EMBL:248457; MID:950392; PID:9563231; MIMD:Y11150

C:Accession: A57285

C:Reference number: A57285; MIMD:952636; PMID:7744966

C:Accession: A57285

C:Molecule type: DNA

C:Residues: 1-1101, 1103-1460 <GNP>

C:Cross-references: GB:L40634

C:Accession: S48457

C:Status: preliminary

C:Molecule type: mRNA

C:Residues: 1-263 <TH>

C:Cross-references: GB:568799; MID:954544; PID:9545450

C:Genetics: sna-1

C:Gene: sna-1

C:Query Match

Best Local Similarity 86.1% Score 31; DB 2; Length 1460;
Matches 5: Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYFITS 7
|||||

b 1190 KEYFITS 1196

C:Species: Brachydanio rerio (zebra fish)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #extl_change 31-Jan-1997

C:Accession: J51225

C:Thisse, C.; Thisse, B.; Schilling, T.F.; Postlethwait, J.H.

C:Development 119, 1203-1215, 1993

C:Title: Structure of the zebrafish sna-1 gene and its expression in wild-type, spad

C:Reference number: J51225; MIMD:941954; PMID:840683

C:Accession: J51225

C:Status: preliminary; translated from GB/EMBL/DBJ

C:Molecule type: mRNA

C:Residues: 1-263 <TH>

C:Cross-references: GB:568799; MID:954544; PID:9545450

C:Genetics: sna-1

C:Gene: sna-1

C:Query Match

Best Local Similarity 86.1% Score 31; DB 2; Length 263;
Matches 5: Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYFITS 7
|||||

b 7 VKEYFITS 13

RESULT 11

phosphotransferase system enzyme II, galactose specific, protein A - Escherichia coli

C:Species: Escherichia coli

C:Date: 29-Oct-1995 #sequence_revision 29-Oct-1999

C:Accession: S55903

C:Reference number: S55903; MIMD:952636; PMID:7744966

C:Accession: S55903

C:Molecule type: DNA

C:Residues: 1-1101, 1103-1460 <GNP>

C:Cross-references: GB:L40634

C:Accession: S48457

C:Status: preliminary

C:Molecule type: mRNA

C:Residues: 1-263 <TH>

C:Cross-references: GB:568799; MID:954544; PID:9545450

C:Genetics: sna-1

C:Gene: sna-1

C:Query Match

Best Local Similarity 86.1% Score 31; DB 2; Length 1460;
Matches 5: Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 KEYFITS 7
|||||

b 1190 KEYFITS 1196

RESULT 11

phosphotransferase system enzyme II, galactose specific, protein A - Escherichia coli

C:Species: Escherichia coli

C:Date: 29-Oct-1995 #sequence_revision 29-Oct-1999

C:Accession: S55903

C:Reference number: S55903; MIMD:952636; PMID:7744966

Reference number: S55901; MUID:95290497; PMID:7772602
 Accession: S55903
 Status: Preliminary
 Molecule type: DNA
 Residues: 1-150 <NR>
 Cross-references: EMBL:X78437, NID:9509737, FID:GNA56228.1, PID:4509173
 Genes:
 Gene: gata

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3% Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||
 137 LKEYFT 142

SUIT 12
 4976
 s system, galactitol-specific IIA component - Escherichia coli (strain K-12)

Species: Escherichia coli
 Date: 12-Sep-1997 #sequence-revision 17-Sep-1997 #text-change 01-Mar-2002
 Accession: E64976

Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Peters, N.T.; Burland, V.; Riley, M.; Co-
 .; Rose, D.J.; Mau, B.; Shao, Y.
 lence 277, 1453-1462, 1997

Title: The complete genome sequence of Escherichia coli K-12.
 Reference number: A64720; MUID:97426617; PMID:9278503

Accession: E64976
 Status: Preliminary; nucleic acid sequence not shown; translation not shown

Molecule type: DNA

Residues: 1-150 <BLAT>
 Cross-references: GB:AE000294; GR:060096; NID:91786195; FID:AACT5155.1; PID:91788410.
 Experimental source: strain K-12, substrain MG1655

Genetics:

Gene: gata

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3% Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||
 137 LKEYFT 142

SUIT 13

0991
 s system, galactitol-specific IIA component - Escherichia coli (strain O157:H7, substr

Species: Escherichia coli
 Date: 18-Jul-2001 #sequence-revision 18-Jul-2001 #text-change 27-Nov-2001
 Accession: A90991

Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 JA Res. 8, 11-22, 2001

Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 Reference number: A99629; MUID:21156231; PMID:11258796

Accession: A90991
 Status: Preliminary

Molecule type: DNA

Residues: 1-150 <HAY>
 Cross-references: GB:BA000007; PID:BA836720.1; PID:91362565; GSPDB:GN00154

Experimental source: strain O157:H7, substrain RMD 0509952

Genetics:

Gene: EGS2897

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3% Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||

Db 137 LKEYFT 142

RESULT 14

C85836
 s system, galactitol-specific IIA component - Escherichia coli (strain O157:H7, sub

Species: Escherichia coli
 Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 27-Nov-2001
 Accession: C85836

R. Peters, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Hille, L.; Grothbeck, F.T.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potlowski, K.; Apoda
 Nature 409, 529-533, 2001

Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 Reference number: A85480; MUID:21074935; PMID:11206551

Accession: C85836

Status: Preliminary

Molecule type: DNA

Residues: 1-150 <STO>

Cross-references: GB:AE005174; NID:91786195; PID:AACT5151.1; GSPDB:GN00145; UMGF:

Experimental source: strain O157:H7, substrain EDL933

Genetics:

Gene: gata

Query Match 83.3% Score 30; DB 2; Length 150;
 Best Local Similarity 83.3% Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
 :|||||
 Db 137 LKEYFT 142

RESULT 15

S47625
 aspartate carboxyltransferase (EC 2.1.1.2) - wheat (fragment)

Alternate names: aspartate transcarboxylase
 N:Alternat names: aspartate transcarboxylase

C:Species: Triticum aestivum (common wheat)
 C:Date: 19-Mar-1997 #sequence-revision 29-Aug-1997 #text-change 07-May-1999

C:Accession: S47625
 R:Barlett, T.J.; Albangbee, A.; Bruce, I.J.; Donovan, P.J.; Yon, R.J.
 Biochim. Biophys. Acta 1207, 187-193, 1994

A:Title: Endogenous polypeptide-chain length and partial sequence of aspartate trans

A:Reference number: S47625; MUID:94355374; PMID:8075153

Accession: S47625

Status: not compared with conceptual translation

Molecule type: mRNA

Residues: 1-233 <BAR>

Experimental source: var. Avalon

C:Superfamily: oxalibine carboxyltransferase; aspartate/oxalibine carboxyltransferase

C:Keywords: pyrimidine nucleotide biosynthesis; transferase; homotrimer

E1-223/20main: aspartate/oxalibine carboxyltransferase; homology (fragment) <AT>

Query Match 83.3% Score 30; DB 2; Length 233;
 Best Local Similarity 85.7% Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KEYFTS 7
 :|||||
 Db 115 IKEYFTS 121

Search completed: April 8, 2003, 12:05:13
 Job time: 34.9167 secs

GenCore version 5.1.4_P1_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

Run on: April 8, 2003, 12:01:50 ; Search time 12.8333 seconds
(without alignments)

33.347 Million cell updates/sec

Sequence: 1 KEYFITS 7

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Search: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No | Score | Query Match Length | DB ID | Description |
|-----------|-------|--------------------|-----------------------------|-------------------|
| 1 | 30 | 83.3 | 150 10 US-09-741-669-424 | Sequence 424, App |
| 2 | 29 | 80.7 | 5745 10 US-09-815-242-15618 | Sequence 12610, A |
| 3 | 28 | 77.4 | 141 10 US-09-815-242-11289 | Sequence 11479, A |
| 4 | 28 | 77.4 | 141 10 US-09-815-242-11471 | Sequence 11471, A |
| 5 | 28 | 77.8 | 219 10 US-09-782-874-7 | Sequence 3, Appli |
| 6 | 24 | 77.4 | 224 10 US-09-812-020-359 | Sequence 350, App |
| 7 | 24 | 77.4 | 276 12 US-10-033-078-16 | Sequence 16, Appl |
| 8 | 28 | 77.8 | 377 10 US-09-910-430-27 | Sequence 27, Appl |
| 9 | 28 | 77.8 | 431 9 US-10-081-816-5 | Sequence 5, Appli |
| 10 | 24 | 77.4 | 511 9 US-09-784-026-5506 | Sequence 5, Appl |
| 11 | 28 | 77.8 | 573 10 US-09-815-242-10760 | Sequence 10760, A |
| 12 | 28 | 77.8 | 584 10 US-09-770-940-2 | Sequence 2, Appli |
| 13 | 24 | 77.4 | 724 10 US-09-815-242-5093 | Sequence 5093, A |
| 14 | 28 | 77.8 | 1114 10 US-09-782-874-2 | Sequence 2, Appli |
| 15 | 27 | 75.0 | 32 10 US-09-144-838-14 | Sequence 14, Appl |
| 16 | 27 | 75.0 | 32 10 US-09-144-838-13 | Sequence 13, Appl |
| 17 | 27 | 75.0 | 33 10 US-09-144-838-15 | Sequence 15, Appl |
| 18 | 27 | 75.0 | 34 10 US-09-144-838-16 | Sequence 16, Appl |
| 19 | 27 | 75.0 | 45 12 US-10-001-870-147 | Sequence 147, A |

| | | | | |
|----|----|------|---------------------------|-------------------|
| 20 | 27 | 75.0 | 53 9 US-10-014-452-12 | Sequence 12, Appl |
| 21 | 27 | 75.0 | 60 9 US-09-888-938-5 | Sequence 5, Appli |
| 22 | 27 | 75.0 | 62 9 US-09-809-391-486 | Sequence 486, App |
| 23 | 27 | 75.0 | 66 10 US-09-144-838-35 | Sequence 35, Appl |
| 24 | 27 | 75.0 | 66 10 US-09-144-838-37 | Sequence 37, Appl |
| 25 | 27 | 75.0 | 67 10 US-09-144-838-31 | Sequence 31, Appl |
| 26 | 27 | 75.0 | 67 10 US-09-144-838-33 | Sequence 33, Appl |
| 27 | 27 | 75.0 | 67 10 US-09-144-838-36 | Sequence 36, Appl |
| 28 | 27 | 75.0 | 67 10 US-09-144-838-38 | Sequence 38, Appl |
| 29 | 27 | 75.0 | 67 10 US-09-144-838-39 | Sequence 39, Appl |
| 30 | 27 | 75.0 | 67 10 US-09-144-838-41 | Sequence 41, Appl |
| 31 | 27 | 75.0 | 68 9 US-09-792-794-29 | Sequence 29, Appl |
| 32 | 27 | 75.0 | 68 10 US-09-864-761-40180 | Sequence 40180, A |
| 33 | 27 | 75.0 | 68 10 US-09-144-838-19 | Sequence 19, Appl |
| 34 | 27 | 75.0 | 68 10 US-09-144-838-32 | Sequence 32, Appl |
| 35 | 27 | 75.0 | 68 10 US-09-144-838-34 | Sequence 34, Appl |
| 36 | 27 | 75.0 | 68 10 US-09-144-838-40 | Sequence 40, Appl |
| 37 | 27 | 75.0 | 68 10 US-09-144-838-42 | Sequence 42, Appl |
| 38 | 27 | 75.0 | 68 10 US-09-144-838-43 | Sequence 43, Appl |
| 39 | 27 | 75.0 | 68 10 US-09-144-838-45 | Sequence 45, Appl |
| 40 | 27 | 75.0 | 68 10 US-09-195-457-11 | Sequence 11, Appl |
| 41 | 27 | 75.0 | 69 10 US-09-144-838-44 | Sequence 44, Appl |
| 42 | 27 | 75.0 | 69 10 US-09-144-838-46 | Sequence 46, Appl |
| 43 | 27 | 75.0 | 89 10 US-09-144-477-8 | Sequence 4, Appli |
| 44 | 27 | 75.0 | 91 8 US-09-779-929-21 | Sequence 21, Appl |
| 45 | 27 | 75.0 | 91 9 US-10-017-275-8 | Sequence 8, Appli |

ALIGNMENTS

RESULT 1
US-09-741-669-424
Sequence 424, Application: 15, 0.47069
FASTA ID: US-09-741-669-424
GENERAL INFORMATION:
APPLICANT: Forsyth, P. Allyn
APPLICANT: Ohlsen, Karl L.
TITLE OR INVENTION: Gees identified as required for
FILE REFERENCE: Elifira 009A
CURRENT APPLICATION NUMBER: US-09-741-669-424
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US-60/370055
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 424
LENGTH: 150
TYPE: PRI
ORGANISM: Escherichia coli
US-09-741-669-424

Query Match: 83.3% Score 40; DB ID: Length 150;
Best local similarity at 78; Pred. NO: 61;
Matches: 5; Conservative: 1; Mismatches: 0; Indels: 0; Gaps: 0;

27 1 KEYFIT 6
CP 127 KEYFIT 142

RESULT 2
US-09-815-242-12610
Sequence 12610, Appl: 15, 0.5156 US-09-815-242-12610
PATENT NO: US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Tawick, John D.

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN
RNA-DIRECTED RNA POLYMERASE (RDRP)

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/647,749, 27;
FILING DATE: 09-Feb-2001
CLASSIFICATION: <UNKNOWN>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/841,583
FILING DATE: 05-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F.
REGISTRATION NUMBER: 27,794
REFERENCE/PATENT NUMBER: Mps-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9090
TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDEDNESS: <UNKNOWN>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-782-874-3

Query Match 77.8% Score 28; DB 10; Length 218;
Best Local Similarity 77.1%; Pred No. 2.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 KEYS
144 VEEYFTN 150

RESULT 6
US-09-912-020-350
Sequence 350, Application US/09/0412000
Patent No. US2002004592A1
GENERAL INFORMATION:
APPLICANT: Zyskind, Judith
APPLICANT: Ohlsen, Karl L.
APPLICANT: Trivick, John
APPLICANT: Forsyth, R. Allan
APPLICANT: Froelich, Jamie M.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: GENE IDENTIFIERS AS PEPTIDES FOR DIFFERENTIATION IN
TITLE OF INVENTION: ESCHERICHIA COLI
FILE REFERENCE: ELITPA 001D1V1
CURRENT APPLICATION NUMBER: US/09/0412000
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: 09/432,704
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: 60/117,465
PRIOR FILING DATE: 1999-01-27
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FASTED for Windows Version 3.0

SEQ ID NO 350
LENGTH: 234
TYPE: PRT
ORGANISM: E. COLI
US-09-912-020-350

Query Match 77.8% Score 28; DB 10; Length 234;
Best Local Similarity 83.3%; Pred No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYS
224 IKEYFT 229

RESULT 7
US-10-033-078-16
Sequence 16, Application US/10/033078
Patent No. US2002012100A1
GENERAL INFORMATION:
APPLICANT: Furukuchi, Yasuhiro
APPLICANT: Hoshino, Tatsuo
APPLICANT: Kimura, Hiroshi
APPLICANT: Kiyasu, Tatsuya
APPLICANT: Nagahashi, Yoshie
TITLE OF INVENTION: PROTEIN BIOSYNTHESIS GENES
FILE REFERENCE: Biotin Genes
CURRENT APPLICATION NUMBER: US/10/033078
CURRENT FILING DATE: 2001-12-27
PRIOR APPLICATION NUMBER: 09/594,184
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: EP 961155A0.5
PRIOR FILING DATE: 1996-09-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver 2.1
SEQ ID NO 16
LENGTH: 276
TYPE: PRT
ORGANISM: Kurthia sp.
US-10-033-078-16

Query Match 77.8% Score 28; DB 10; Length 276;
Best Local Similarity 57.1%; Pred No. 2.7e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 KEYS
214 VEEYFTS 220

RESULT 8
US-09-913-430-27
Sequence 27, Application US/09/0910430
Patent No. US2002012735A1
GENERAL INFORMATION:
APPLICANT: Godtrot, Edmond
APPLICANT: Bollen, Alex
APPLICANT: Lebouille, Gerard
TITLE OF INVENTION: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF
TITLE OF INVENTION: PROTEINS, EXPRESSED IN THE XODUS PLOINUS SALIVARY
FILE REFERENCE: VANM229.001CP1
CURRENT APPLICATION NUMBER: US/09/0910430
CURRENT FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: PCT/FR99/0061
PRIOR FILING DATE: 1999-06-06
PRIOR APPLICATION NUMBER: GB 991342.5
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FASTED for Windows Version 3.0
SEQ ID NO 27
LENGTH: 377
TYPE: PRT

ORGANISM: Ixodes ricinus
US-09-910-430-27

Query Match 77.8% Score 28; DB 10; Length 377;
Best Local Similarity 57.1% Pred No. 3.6e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 KEYEFTS 7
1:11111
290 VKDFTS 296

RESULT 9
US-10-081-816-5
Sequence 5; Application US/0908165
Publication No. US20030045472A1

GENERAL INFORMATION:

APPLICANT: Axel Richard

APPLICANT: Scott, Kristin

TITLE OF INVENTION: Chemosensory Gene Family Encoding Gustatory And Olfactory Receptors

FILE REFERENCE: 0575-6474-A/PM/AM

CURRENT APPLICATION NUMBER: US/10/081.816

PRIOR FILING DATE: 2002-02-22

PRIOR APPLICATION NUMBER: 60/271,319

NUMBER OF SEQ ID NOS: 116

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 5

LENGTH: 431

TYPE: PRT

ORGANISM: Drosophila melanogaster

US-10-081-816-5

Query Match 77.8% Score 28; DB 9; Length 421;
Best Local Similarity 43.3% Pred No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYEFT 6
1:11111
286 IEVEY 291

RESULT 10
US-09-748-626-5606
Sequence 5606; Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NACKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/748.626

PRIOR FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-18

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO: 5606

LENGTH: 511

TYPE: PRT

ORGANISM: Corynebacterium glutamicum
US-09-738-626-5606

Query Match 77.8% Score 48; DB 9; Length 511;
Best Local Similarity 83.3% Pred No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEYEFT 6
1:11111
DB 386 IEVEY 391

RESULT 11
US-09-815-242-10760
Sequence 10760; Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl D.

APPLICANT: Tyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: FLITRA.011A

CURRENT APPLICATION NUMBER: US/09/815.242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

FILE REFERENCE: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 10760

LENGTH: 570

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-815-242-10760

Query Match 77.8% Score 48; DB 10; Length 570;
Best Local Similarity 100.0% Pred No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEYEFT 6
1:11111
DB 247 KEVEY 251

RESULT 12
US-09-770-940-2
Sequence 2; Application US/09770940
Patent No. US20020009790A1

GENERAL INFORMATION:

APPLICANT: Christensen, T.

APPLICANT: Thorsen, H.

APPLICANT: Kreider, J.

APPLICANT: Buchholt, H.

TITLE OF INVENTION: PROCESS FOR STABILIZING PROTEINS IN AN

FILE REFERENCE: DYO014.001APC

```

CURRENT APPLICATION NUMBER: US/09/770,940
CURRENT FILING DATE: 2001-01-24
EARLIER APPLICATION NUMBER: 08/983,364
EARLIER FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 584
TYPE: PRT
ORGANISM: Citrus hawaiiensis var. class 1
US-09-770-940-2

Query Match:
Best Local Similarity: 77.8%; Score 28; DB 10; Length 584;
Pred. No. 5.4e+02;
Matches: 5; Conservative: 1; Mismatches: 0; Indels: 0; Gaps: 0;

Y 1 KEYFTS 7
  11111
Db 4 KEYFTS 9

RESULT 13
US-09-815-242-5093
Sequence 5093: Application US/09/815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Ohlssen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITPA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5093
LENGTH: 723
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5093

Query Match:
Best Local Similarity: 77.8%; Score 28; DB 10; Length 723;
Pred. No. 6.6e+02;
Matches: 4; Conservative: 3; Mismatches: 0; Indels: 0; Gaps: 0;

Y 1 KEYFTS 7
  11111
Db 275 VREFTS 281

RESULT 14
US-09-782-874-2
Sequence 2: Application US/09/782874

```

```

Patent No. US20010023067A1
GENERAL INFORMATION:
APPLICANT: Wassenecker, Michael
Riedel, Leonhard
Schiebel, Manfred
Sanger, Heinz
TITLE OF INVENTION: NOVEL ACT. MOLECULES ENCODING
POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN
RNA-DIRECTED RNA POLYMERASE (RDPR)
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-647,874
FILING DATE: 08-Feb-2001
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/411,583
FILING DATE: 05-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Haley, James F.
REGISTRATION NUMBER: 27,764
REFERENCE/DECKET NUMBER: m-3-1
TELECOMMUNICATION INFORMATION
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1114 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-782-874-2

Query Match:
Best Local Similarity: 77.8%; Score 28; DB 10; Length 1114;
Pred. No. 9.9e+02;
Matches: 4; Conservative: 3; Mismatches: 0; Indels: 0; Gaps: 0;

Y 1 KEYFTS 7
  11111
Db 843 VEEFTS 849

RESULT 15
US-09-144-838-14
Sequence 14: Application US/09/144838A
Patent No. US2002005196A1
GENERAL INFORMATION:
APPLICANT: Siani, Michael A.
APPLICANT: Wilken, Jill
APPLICANT: Simon, Reyna
APPLICANT: Kent, Stephen B.H.
TITLE OF INVENTION: Modular Protein Libraries and Methods of Preparation
FILE REFERENCE: GFPN-020/0105
CURRENT APPLICATION NUMBER: US/09/144,838A
CURRENT FILING DATE: 1998-08-31
EARLIER APPLICATION NUMBER: US 60/077,620
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 14
LENGTH: 32

```

TYPE: PRI
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
-09-144-838-14

Query Match 75.0% Score 27; DB 10; Length 32;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IKEYF 5
|||||
23 IKEYF 27

arch completed: April 8, 2003, 12:06:21
b time : 13.8333 secs

 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 C -----
 R EMBL: A6001550; AAD06952.1; -
 R InterPro: IPR004552; AGP_Acyltrn.
 R InterPro: IPR002123; Acyltransferase.
 R Pfam: PF01553; Acyltransferase; 1.
 R TIGfam: TIGR00530; AGP_Acyltrn; 1.
 W Phospholipid biosynthesis; Transferase; Acyltransferase;
 W Inner membrane; Complete proteome.
 Q SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CPG64;
 Query Match 100.0%; Score 25; DB 1; Length 237;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y 1 LESYT 5
 |||||
 b 204 LESYT 208

 RESULT 4
 LSC_HELPV STANDARD; PRT; 240 AA.
 C 025603;
 T 15-DEC-1998 (Rel. 37, Created)
 T 15-DEC-1998 (Rel. 37, Last sequence update)
 T 15-JUN-2002 (Rel. 41, Last annotation update)
 T 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
 E acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
 (ADPAT)
 N PISC OR HP1348.
 C Helicobacter pylori (Campylobacter pylori).
 C Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 C Helicobacter.
 X NCBI_TaxID=2190;
 N [1]
 P SEQUENCE FROM N.A.
 C STRAIN=26695; ATCC: 49239;
 X MEDLINE=97394467; PubMed=9252185;
 A Tomb J.-F., White O., Kellavag A.R., Clayton R.A., Sutton G.G.,
 A Fleischmann P.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 A Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 A Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 A McKenney K., Fitzgerald L.W., Lee N., Adams M.D., Hickey E.K.,
 A Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 A Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
 A Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 A Venter J.C.;
 T "The complete genome sequence of the gastric pathogen Helicobacter
 T pylori."
 L Nature 388:539-547(1997).
 C ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
 C ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
 C -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -
 C CoA + 1,2-diacyl-sn-glycerol 3-phosphate
 C -1- FUNCTION: De novo phospholipid biosynthesis; second step.
 C -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
 C -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
 C ACYLTRANSFERASE FAMILY.
 C -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

 CC or send an email to license@isb-sib.ch).
 CC -----
 R EMBL: A600636; AAD08393.1; -
 R TIGR: HP1348; -
 R InterPro: IPR004552; AGP_Acyltrn.
 R InterPro: IPR002123; Acyltransferase.
 R Pfam: PF01553; Acyltransferase; 1.
 R TIGfam: TIGR00530; AGP_Acyltrn; 1.
 W Phospholipid biosynthesis; Transferase; Acyltransferase;
 W Inner membrane; Complete proteome.
 Q SEQUENCE 240 AA; 27745 MW; 22B1D0EB190BDD CPG64;
 Query Match 100.0%; Score 25; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Y 1 LESYT 5
 |||||
 b 204 LESYT 208

 RESULT 5
 Y711_METUA STANDARD; PRT; 322 AA.
 ID Y711_METUA
 AC 058121;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein M0711.
 GN M0711.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 RX NCBI_TaxID=2190;
 N [1]
 P SEQUENCE FROM N.A.
 C STRAIN=JAL-1 / DSM 2661 / ATCC 43047;
 RC MEDLINE=96337999; PubMed=8688087;
 RX Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kellavag A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.L.,
 RA Overbeek P., Kirkness E.F., Weinstock K.G., Merriam J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Patton M.E., Roberts J.M., Hurst M.A., Yano R.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii."
 PI Science 273:1058-1073(1996).
 CC -----
 C This SWISS-PROT entry is copyright. It is produced through a collaboration
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -
 C the European Bioinformatics Institute. There are no restrictions on its
 C use by non-profit institutions as long as its content is in no way
 C modified and this statement is not removed. Usage by and for commercial
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C or send an email to license@isb-sib.ch).
 C -----
 R EMBL: U67518; AAB98714.1; -
 R TIGR: M0711; -
 R Hypothetical protein; Transmembrane; Complete proteome.
 R TRANSMEM 239 319 POTENTIAL.
 Q SEQUENCE 322 AA; 38631 MW; 07C14417504927 CPG64;
 Query Match 100.0%; Score 25; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Y 1 LESYT 5
 |||||
 b 106 LESYT 110

 Query Match 100.0%; Score 25; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

of release factors".

Mol. Biol. Evol. 17:982-989(2000).

1- FUNCTION: Directs the termination of nascent peptide synthesis (translation) in response to the termination codons UAA, UAG and UGA (by similarity).

1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, ONE OF WHICH BINDS GTP.

1- SUBCELLULAR LOCATION: CYTOSOL; ER (by similarity).

1- SIMILARITY: BELONGS TO THE EUKARYOTIC RELEASE FACTOR 1 FAMILY

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre> or send an email to license@ebi.ac.uk).

EMBL: AF198107; AAF74402.1

HSSP: P46055; 1079.

InterPro: IPR004403; eRF1.

InterPro: IPR005140; eRF1.

InterPro: IPR005142; eRF1.

PIfam: PF03463; eRF1_1; 1.

PIfam: PF03464; eRF1_2; 1.

PIfam: PF03465; eRF1_3; 1.

TIGRFAMs: TIGR00108; eRF1; 1.

Protein biosynthesis.

SEQUENCE 457 AA; 51029 MW; EE185FFD1F0C943 CRC64;

Query Match 100.0%; Score 25; DB 1; Length 457;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LESYT 5

|||||

30 LESYT 34

RESULT 9

RXRG_XENLA STANDARD; PRT; 470 AA.

AC P51129;

DT 01-OCT-1996 (Rel. 34; Last sequence update)

DT 01-OCT-1996 (Rel. 34; Last sequence update)

DT 15-MAR-1999 (Rel. 38; Last annotation update)

DE Retinoic acid receptor RXR-gamma.

DE RXRG OR NR2B3.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopus; Xenopus.

OX NCBI_TaxID=8355;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE:9236110; PubMed:112717;

RA Blumberg P., Mangelsdorf B.J., Dyck J.A., Blumberg P.A., Evans R.M., de Robertis E.M.,

RT "Multiple retinoid-responsive receptors in a single cell: families of retinoid 'X' receptors and retinoid acid receptors in the Xenopus egg."

RU Proc. Natl. Acad. Sci. U.S.A. 89:4341-4345(1992).

CC 1- FUNCTION: INVOLVED IN RETINOIC ACID RESPONSE PATHWAY. BINDS 9-CIS RETINOIC ACID (9CRA) (BY SIMILARITY).

CC 1- SUBUNIT: HOMODIMER OR HETERODIMER (BY SIMILARITY).

CC 1- SUBCELLULAR LOCATION: Nuclear.

CC 1- DEVELOPMENTAL STAGE: IT IS SYNTHESIZED DURING OOGENESIS AND PERSISTS IN THE CLEAVING EMBRYO AT APPROXIMATELY CONSTANT LEVELS UNTIL IT IS DEGRADED JUST BEFORE GASTRULATION.

CC 1- DOMAIN: COMPOSED OF THREE DOMAINS, A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

CC 1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC NR2 SUBFAMILY.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre> or send an email to license@ebi.ac.uk).

EMBL: U1143; NOT_ANNOTATED_CDS

HSSP: P19793; 2NLL.

TRANSFAC: T01960;

InterPro: IPR000536; Hormone_rec_194.

InterPro: IPR001723; Steroid_receptor.

InterPro: IPR001729; Znf_C4steroid

PIfam: PF00104; hormone_rec; 1.

PIfam: PF00105; Zf-C4; 1.

PRINTS: PR00398; STEROIDOMER.

PRINTS: PR00047; STEROIDOMER.

ProDom: PD00035; Znf_C4steroid; 1

SMART: SM00430; HOL1; 1.

SMART: SM00399; Znf_C4; 1.

PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.

DR PROSITE: Transcription regulation: DNA-binding; Nuclear protein; Zinc-finger; Multigene family.

KW ZINC-FINGER; MULTIGENE FAMILY.

FT 1 145 MODULATING (BY SIMILARITY).

FT 2 211 NUCLEAR RECEPTOR-TYPE.

FT 3 146 166 C4-TYPE.

FT 4 182 205 C4-TYPE.

FT 5 212 235 HINGE.

FT 6 470 LIGAND BINDING (BY SIMILARITY).

FT 7 236 470

SEQUENCE 470 AA; 51925 MW; DD1172D53CF3A5F CRC64;

Query Match 100.0%; Score 25; DB 1; Length 470;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LESYT 5

|||||

408 LESYT 412

RESULT 10

MURD_TREPA STANDARD; PRT; 532 AA

AC O83873;

DT 15-DEC-1998 (Rel. 37; Last sequence update)

DT 15-DEC-1998 (Rel. 37; Last sequence update)

DT 15-JUN-2002 (Rel. 41; Last annotation update)

DE UDP-N-acetylaminotyrosyl-L-alanyl-D-glutamate synthetase (D-glutamate acid adding enzyme).

DE MURD OR TP0903.

OS Treponema pallidum.

OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_TaxID=160;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE:9642770; PubMed:9665876;

RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Oakes R., Gwin M., Hickey E.K., Clayton P., Potluri K.A., Solerger E., Hardham T.M., Melton M.P., Salzberg S., Peterson J., Khalay H., Richardson D., Howell I.K., Chidambaram M., Osterback T., McDonald L., Attlich P., Bowman G., Cotton M.D., Fujii C., Gailand S., Hitch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.,

RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete."

RU Science 291:375-382(1998).

CC 1- FUNCTION: CELL WALL FORMATION. CATALYZES THE ADDITION OF D-GLUTAMATE TO THE NUCLEOTIDE PRECURSOR UDP-N-ACETYLAMINOYL-L-ALANINE (UMA) (BY SIMILARITY).

```

1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmutamoyl-L-alanine + D-
  glutamate -> ADP + phosphate + UDP-N-acetylmutamoyl-L-alanyl-D-
  glutamate
1- PATHWAY: Peptidoglycan biosynthesis
1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity)
1- SIMILARITY: BELONGS TO THE MURCEDEF FAMILY
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
EMBL: AE001259; AAC65856.1; -
HSSP: P14900; IEH.
R TIGR: TP0903; -
R InterPro: IPR000713; Mur_Ligase.
R InterPro: IPR004101; Mur_Ligase_C.
R Pfam: PF01225; Mur_Ligase_2.
R Pfam: PF02875; Mur_Ligase_C1.
R TIGRPFAM: TIGR01087; MurL1.
R Peptidoglycan synthetase; Cell wall; Cell division; Ligase;
  ATP-binding; Complete proteome.
  NP_BIND: 124 130
  1 SEQUENCE 532 AA: 57847 MW: 97035.46476646 PKC64
Query Match 100.0% Score 25; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LESYT 5
|||||
385 LESYT 389
-----
RESULT 11
370_CHICK STANDARD: PRT: 634 AA.
1- AUG-1988 (Rel. 08, Created)
1- AUG-1988 (Rel. 08, Last sequence update)
1- FEB-1994 (Rel. 28, Last annotation update)
Heat shock 70 kDa protein (HSP70).
Gallus gallus (Chicken)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
NCBI_TaxID: 9031.
[1]
SEQUENCE FROM N.A.
MEDLINE: 86304452; PubMed: 3017985;
Mortimore R.L., Hunt C., Huang S.Y., Berg K.L., Banerji S.S.;
"Organization, nucleotide sequence, and transcription of the chicken
HSP70 gene."
J. Biol. Chem. 261:12692-12699(1986).
1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
EMBL: J02579; AAA48825.1; -
PIR: A25646; A25646.
HSSP: P08109; ICKR
InterPro: IPR001023; HSP70.
Pfam: PF00012; HSP70_1
PRINTS: PR00301; HEATSHOCK70.

```

```

DR PROSITE: PS00297; HSP70_1.
DR PROSITE: PS00329; HSP70_2.
DR PROSITE: PS01036; HSP70_3.
KW ATP-binding; Heat shock; Multigene family.
SQ SEQUENCE 634 AA: 69750 MW: 42706.7708365AEB PKC64
Query Match 100.0% Score 25; DB 1; Length 634;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LESYT 5
|||||
545 LESYT 549
-----
RESULT 12
HS72_HUMAN STANDARD: PRT: 639 AA.
ID HS72_HUMAN
AC P54552; O15508; Q9UE78;
RT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-OWN-2002 (Rel. 42, Last annotation update)
DE Heat shock-related 70 kDa protein 2 (Heat shock 70 kDa protein 2);
  HSPA2.
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID 9606;
RN [1]
RX MEDLINE: 96130116; PubMed: 7829106;
RA Bonycastle L.L.C., Yu C.-E., Hunt C.R., Trask B.J., Clancy K.P.,
  Weber J.L., Patterson D., Schellenberg G.D.;
  "Cloning, sequencing, and mapping of the human chromosome 14 heat
  shock protein gene (HSPA2)."
  J. Biol. Chem. 269:2385-2393(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Goralski T.J., Krensky A.M.;
  Submitted (Apr-1996) to the EMBL/Genbank/CCDB databases
RN [3]
RX MEDLINE: 95152405; PubMed: 7849706;
RA Poux A.-F., Nguyen V.T., Squire J.A., Cox D.W.;
  "A heat shock gene at 14q32: mapped and expression."
  Hum. Mol. Genet. 3:1819-1822(1994).
OC PREEXISTENT PROTEINS AGAINST AGGREGATION AND MEDATE THE FOLDING
  OF NEWLY TRANSLATED POLYPEPTIDES IN THE CYTOSOL AS WELL AS WITHIN
  ORGANELLES. THESE CHAPERONES PARTICIPATE IN ALL THESE PROCESSES
  THROUGH THEIR ABILITY TO RECOGNIZE NONNATIVE CONFORMATIONS OF
  OTHER PROTEINS. THEY BIND EXTENDED PEPTIDE SEGMENTS WITH A NET
  HYDROPHOBIC CHARACTER EXPOSED BY POLYPEPTIDES DURING TRANSLATION
  AND MEMBRANE TRANSLOCATION, OR FOLLOWING STRESS-INDUCED DAMAGE.
1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
EMBL: U26396; AAA52698.1; -
EMBL: U56725; AAD11466.1; -
EMBL: U01049; AAC50076.1; -
HSSP: P19120; HSC.
DR Genes: HSCA-5235; HSPA2.
DR MIM: 140560; -
InterPro: IPR001023; HSP70.
DR Pfam: PF00012; HSP70_1.

```



```

InterPro: PF005259: PriA.
Pfam: PF00271: helicase_C. 1.
SMART: SM00487: DEXDC_1.
TIGRFAMs: TIGR00595: PriA_1.
DNA replication: DNA-binding: ATP-binding: Helicase: Primosome:
Zinc-finger: Complete proteome.
NP_BIND: 241 248 ATP (POTENTIAL).
SITE 337 340 DEH BOX.
2N_FING 458 470 C4-TYPE (POTENTIAL).
ZN_FING 485 501 C4-TYPE (POTENTIAL).
SEQUENCE 753 AA: 64764 MW: 674577 PC: AB05 C: C066:

Query Match 100.0% Score 25; DB 1; Length 753;
Best Local Similarity 100.0% pred. No. 1; e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Query Match      100.0%; Score 25; DB 1; Length 753;
Best Local Similarity 100.0%; Fred. No. 1-le-02;
Matches      5; Conservative      0; Indels      0; Gaps
QY      1 LESYT 5
        |||||
Db       379 LESYT 383

Search completed: April 8, 2003, 12:02:14
Job time : 15.0833 secs

```

```

1 LESYT 5
  |||||
379 LESYT 383

```

| SOLT 15 | IA_CHLTR | STANDARD: | PRT: | 753 AA |
|---------|------------|-----------|------|--------|
| | PRTA_CHLTR | | | |

New peptides useful for inhibiting human immuno-deficiency virus type 1 (HIV-1) gp120 induced neuronal cell death

Claim 1: Page 11: 16pp: English.

AAV46695 and AAV46696 represent peptides which inhibit HIV-1 gp120 induced neuronal cell death. Pharmaceutical compositions containing the peptides are useful for treating symptoms caused by neuronal cell loss. Such conditions especially associated with HIV infection include encephalopathies, neuropathies, memory loss, dementia, depression, psychosis and opportunistic infections. The peptides act as antagonists of gp120-mediated neurotoxicity and subsequent neuronal degeneration. This enables therapeutic treatment of HIV infection and other inflammatory neurological diseases, including multiple sclerosis, tropical spastic paraparesis and Alzheimer's disease.

Sequence 7 AA:

Query Match 100.0% Score 16; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 7; Ho-05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IKEYPTS 7
|||||||
1 IKEYPTS 7

JULI 2
38988
ABP38988 standard: Protein: 204 AA.

ABP38988:

24-JUL-2002 (first entry)

Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3833.

Staphylococcus epidermidis: open reading frame: ORF: bacterial infection, antibacterial: gene therapy.

Staphylococcus epidermidis:

US6380370-B1.

30-APR-2002.

13-AUG-1998. 3833 0134001.

14-AUG-1997. 97US-055779P.

08-NOV-1997. 97US-064964P.

(GENO-) GENOME THERAPEUTICS CORP.

Doucette-Stamm LA, Bush D;

WPI: 2002-381255/41.

N-PSDB: ABN81533.

Novel isolated nucleic acid encoding a Staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections.

Disclosure: SEQ ID 3833: 267pp: English.

ABN90538 to ABN93374 represent staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP38124 to ABP37960. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the

CC USPTO web site.
XX
SQ Sequence 204 AA;

Query Match 88.9%; Score 16; DB 23; Length 204;
Best Local Similarity 71.4%; Pred. No. 97;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKEYPTS 7
|||||||
DB 195 VKFFFA 201

RESULT 3
AAW3792
ID AAW3792 standard: Protein: 422 AA.

XX
AC AAW3792;

XX
DT 10-NOV-1997 (first entry)

XX
DE Cytidine monophospho-N-acetylneuraminic acid hydroxylase.

XX
KM CMP-Neu5Ac hydroxylase; CMP-Neu5Ac hydroxylase; enzyme activity;
KW glycoprotein; cytochrome b5; tumour; human; pig.

XX
OS Sus scrofa.

XX
PN W05703200-A1.

XX
PD 30-JAN-1997.

XX
PF 05-JUL-1996. 96WO-EP02966.

XX
PR 22-SEP-1995. 95EP-0114930.

XX
PR 07-JUL-1995. 95EP-0110609.

XX
PA (POFF) ROEHRIGER MANNHEIM GMBH.

XX
PI Haselbeck A, Honold K, Kelm S, Schaner R, Schlenzka W;

XX
PI Shaw L;

XX
DR WPI: 1997-119055/11.

XX
DR N-PSDB: AAT78596.

XX
PT New CMP-N-acetylneuraminic acid hydroxylase and related DNA - used
PI esp. to detect enzyme activity in mammalian cells, allowing prodn.
PI of glyco(protein)s without N-glycosyl-neuraminic acid suitable for
PI human use

XX
PS Claim 2: Page 32-33; 44pp: English.

XX
XX The present sequence represents a now isolated protein cytidine
CC monophospho-N-acetylneuraminic acid hydroxylase, which is uncontaminated
CC by cytochrome b5 or cytochrome b5 reductase. The protein is produced by
CC exogenous DNA in prokaryotic or eukaryotic cells and it catalyses
CC conversion of CMP-Neu5Ac to CMP-Neu5Gc (Neu5Ac and Neu5Gc - N-acetyl or
CC glycosyl-neuraminic acid respectively). Oligonucleotides which hybridise
CC to DNA encoding the protein are used to detect such nucleic acids by
CC hybridisation assays, and are used to inhibit expression of the protein,
CC i.e. as antisense molecules for treatment of tumours, particularly those
CC of high metastatic potential. Cells that have been modified so that they
CC no longer produce the protein are used to produce Neu5Gc-free
CC glyconjugates for therapeutic use. The DNA encoding the protein is
CC used to produce recombinant protein and detection of the gene, normally
CC dormant, can be used to diagnose tumours. Glycoproteins that lack Neu5Gc
CC more closely resemble human products and are suitable for long term
CC therapy since they contain no non-human epitopes (Neu5Gc containing
CC proteins are immunogenic in humans).

XX
SQ Sequence 422 AA;
Query Match 88.9%; Score 16; DB 18; Length 422;

Best Local Similarity: 100.0%; Pred. No. 1.9e+02;
Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 IKEYFT 6
|||||
303 IKEYFT 308

RESULT 4

AAW10735 standard; Protein: 422 AA.

AAW10735;

30-SEP-1997 (first entry)

Porcine CMP-N-acetyl-neuraminic acid hydroxylase.

N-acetyl-neuraminic acid; Neu5Ac; sialic acid; porcine; CMP; cytidine monophosphate; antisense therapy; recombinant production; N-glycolylneuraminic acid side chain deficient; erythropoietin.

Sus scrofa.

EP752474-A1.

08-JAN-1997.

07-JUL-1995; 95EP-0110609.

07-JUL-1995; 95EP-0110609.

(BOE) BOEHRINGER MANNHEIM GMBH.

Haselebeck A, Honold K, Kelm S, Schauer R, Schlentzka W;

Shaw L;

WPI: 1997-067451/07.

N-PSDB: AAI48588.

Porcine CMP-N-acetyl-neuraminic acid hydroxylase - catalyzes the conversion of CMP-Neu5Ac to CMP-Neu5Gc; antisense sequences of DNA encoding CMP-N-acetyl-neuraminic acid hydroxylase are used for prodn. of modified glycoproteins

Claim 2; Page 14-19; 24pp; English.

AAW10735 represents a porcine cytidine mono phosphate (CMP-N-acetyl-neuraminic acid (CMP-Neu5Ac) hydroxylase enzyme. The enzyme catalyzes the conversion of CMP-Neu5Ac to CMP-Neu5Gc (cytidine monophosphate-N-glycolylneuraminic acid). Antisense sequences of this cDNA sequence are used for the production of CMP-Neu5Gc hydroxylase-deficient mammalian cells. The modified cells are useful for the production of recombinant glycoproteins lacking N-glycolylneuraminic acid side chains, e.g. erythropoietin.

Sequence 422 AA;

Query Match: 88.9%; Score 32; DB 18; Length 422;

Best Local Similarity: 100.0%; Pred. No. 1.9e+02;

Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 IKEYFT 6
|||||
303 IKEYFT 308

RESULT 5

ABP26246 standard; Protein: 728 AA.

ABP26246;

02-JUL-2002 (first entry)

Streptococcus polypeptide SEQ ID NO: 1668.

Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae; group A streptococcus; Streptococcus pyogenes; antibacterial; antiinflammatory; infection; vaccine; meningitis; gene therapy.

Streptococcus pyogenes

W0200234771-A2.

02-MAY-2002.

27-OCT-2000; 2000GB-0026333.

24-NOV-2000; 2000GB-0028727.

07-MAR-2001; 2001GB-0005640.

(CHIR-) CHIRON SPA.

(GENO-) INST GENOMIC RES.

Telford J, Masignani V, Margarit Ros YL, Grandi G, Fraser C;

Tectelin H;

WPI: 2002-352536/38.

N-PSDB: ABN66877.

Claim 1; Page 3318; 4525pp; English.

The invention relates to a protein (ABP26246) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and antibodies that bind (1) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (1) are used to detect Streptococcus in a biological sample. (1) is used to determine whether a compound binds to (1). A composition comprising (1) or a nucleic acid encoding (1), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (1) may be used to recombinantly produce (1) and may be used in gene therapy. Antibodies to (1) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.

Sequence 728 AA;

Query Match: 88.9%; Score 32; DB 23; Length 728;

Best Local Similarity: 71.4%; Pred. No. 3.3e+02;

Matches: 5; Conservative: 2; Mismatches: 0; Indels: 0; Gaps: 0;

1 IKEYFTS 7
|||||
131 VKDYFTS 137

DB

AAW10745 standard; Protein: 1460 AA.

AAW10745;

27-JUL-2001 (first entry)

S cerevisiae apoptosis associated protein YJL12W.

Yeast; fungus; apoptosis; infection; proliferative disease;
vaccine; autoimmune disease; ischaemia; neurodegeneration;

Saccharomyces cerevisiae.

WO200102550-A2.

11-JAN-2001.

03-JUL-2000; JPO/WHO-HE-00077.

01-JUL-1999; 9SEP-0870141.

(JANC) JANSSEN PHARM NV.

Contreras RH, De Backer MD, Luyten WHML, Malcorps IRL,
Nellissen BWM, Reekmans RJ;

WPI: 2001-367042/96.

N-PSDB: AAH29781.

Yeast and fungal nucleic acids encoding proteins involved in a pathway
leading to programmed cell death, useful for treating proliferative
disorders, yeast and fungal infections, or for preventing apoptosis in
certain diseases -

Claim 1: Fig 1: 218pp: English.

The present invention provides the protein and coding sequences of a
number of apoptosis associated proteins from the yeast *Saccharomyces*
cerevisiae and the fungus *Candida albicans*. These can be used to identify
treatments for fungal and yeast infections, for proliferative diseases
and for apoptosis related diseases such as autoimmune diseases, ischaemia
and neurodegeneration. The present sequence is one of the *S. cerevisiae*
proteins of the invention.

Sequence 1460 AA:

Query Match 86.1% Score 31; DB 22; Length 1460;
Best Local Similarity 71.4% Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7

1111111

1190 LKEYPTS 1196

SULT 7

129339

AA029339 standard; Protein: 150 AA.

AA029339;

18-DEC-2001 (first entry)

Novel mar regulated protein (NIMR) #11.

mar regulated polypeptide; NIMR; microbial infection; antibacterial.

Escherichia coli.

WO200107076-A2.

27-SEP-2001.

08-MAR-2001; 2001WO-US07478.

10-MAR-2000; 2000US-188362P.

(TUFT) TUFTS COLLEGE.

Levy SB, Balbosa TM, Alekshun MN;

XX WPI: 2001-603769/68.
XX N-PSDB: AAS46240.

XX Identifying compounds that modulate a newly identified mar regulated
XX polypeptide activity, useful as antimicrobial compounds. Involves
XX contacting the polypeptide with a test compound -

PS Disclosure: Page 202-203; 526pp: English.

XX The invention relates to a method of identifying compounds that modulate
XX a newly identified mar regulated (NIMR) polypeptide activity. The method
XX comprises contacting an NIMR polypeptide with a test compound under
XX interaction conditions, determining the ability of the compound to
XX modulate the activity or expression of the polypeptide, and selecting the
XX modulators. NIMR nucleic acids and polypeptides are used in the treatment
XX of microbial infections, and in screening for modulators of NIMR
XX expression and activity. These modulators can be used to reduce the
XX infectivity of a microbe on a surface, and the virulence of a microbe in
XX a subject suffering from an infection. AA029339-AA029379 represent
XX *Escherichia coli* NIMR amino acid sequences of the invention.

SQ Sequence 150 AA:

Query Match 83.3% Score 31; DB 22; Length 150;
Best Local Similarity 83.3% Pred. No. 1.0e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KEYPT 6

DB 137 LKEYPT 142

RESULT 8
AA098376
ID AA098376 standard; Protein: 150 AA.

XX AA098376;

XX 21-SEP-2001 (first entry)

XX *Escherichia coli* protein sequence SW: ID NO:424.

XX *Escherichia coli*; identification; proliferation; microorganism;
XX antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
XX bacterial growth inhibition.

XX *Escherichia coli*.

XX WO200148209-A2.

XX 05-JUL-2001.

XX 19-DEC-2000; 2000WO-US34419.

XX 23-DEC-1999; 99US-0173005.

XX (ELIT-) ELITRA PHARM INC.

XX Forsyth PA, Ohlsen KL, Zyskind JM.

XX WPI: 2001-457376/49.

XX N-PSDB: AAH81432.

XX Novel nucleic acids encoding proteins required for *Escherichia coli*
XX proliferation, useful for screening for antimicrobial agents -

PS Claim 19; Page 544; 596pp; English

XX The present invention describes a purified or isolated nucleic acid
XX sequence (I) consisting essentially of one of the 93 nucleotide sequences
XX given in AAH81202 to AAH81294, wherein expression of the nucleic acid in a
XX microorganism is capable of inhibiting proliferation of a microorganism.

(1) have antibacterial and antibiotic activities, and can be used in gene therapy. Expression of (1) in a microorganism inhibits proliferation of the microorganism, and the manufactured antibiotic is useful for reducing the activity or level of a gene product required for proliferation of a microorganism in a subject, specifically humans. The nucleic acids that inhibit bacterial growth or proliferation can be used as antisense therapeutics for killing bacteria. In addition to therapeutic applications, the nucleic acid sequences complementary to sequences required for proliferation can be used as diagnostic tools. For example, nucleic acid probes complementary to proliferation-regulated sequences that are specific for particular species of microorganisms can be used as probes to identify particular microorganism species in clinical specimens. AAH81246 to AAH1487 encode the Escherichia coli proteins given in AAG94236 to AAG94431, and AAH81494 to AAH81491 represent oligonucleotides, which are used in the exemplification of the present invention.

Sequence 150 AA:

Query Match 81.3% Score 30: DB 22: Length 150:

Best Local Similarity 81.3% Frq# No. 1.8e+02:

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

1 IKERYT 6

137 IKERYT 142

ESOUT 9

AG42212

AG42212 standard: Protein: 258 AA

AG42212:

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO. 52616.

Protein identification: signal transduction pathway: metabolic pathway:

hybridisation assay: genetic mapping: gene expression control: promoter:

termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000: 0303439.

25-FEB-1999: 99US-0131825.

05-MAR-1999: 99US-0131810.

09-MAR-1999: 99US-0132548.

23-MAR-1999: 99US-0125788.

25-MAR-1999: 99US-0126264.

29-MAR-1999: 99US-0126785.

01-APR-1999: 99US-0127462.

06-APR-1999: 99US-0128234.

08-APR-1999: 99US-0128714.

16-APR-1999: 99US-0129845.

19-APR-1999: 99US-0130077.

21-APR-1999: 99US-0130449.

23-APR-1999: 99US-0130510.

28-APR-1999: 99US-0131449.

30-APR-1999: 99US-0132048.

04-MAY-1999: 99US-0132407.

05-MAY-1999: 99US-0132484.

06-MAY-1999: 99US-0132485.

07-MAY-1999: 99US-0132486.

11-MAY-1999: 99US-0132487.

99US-0132563.

99US-0134256.

99US-0134218.

99US-0134219.

99US-0134221.

99US-0134370.

99US-0134768.

99US-0134941.

99US-0135124.

99US-0135353.

99US-0135629.

99US-0136021.

99US-0136392.

99US-0136782.

99US-0137222.

99US-0137528.

99US-0137502.

99US-0137724.

99US-0138094.

99US-0138540.

99US-0138847.

99US-0139419.

99US-0139452.

99US-0139453.

99US-0139454.

99US-0139455.

99US-0139456.

99US-0139457.

99US-0139458.

99US-0139459.

99US-0139460.

99US-0139461.

99US-0139462.

99US-0139463.

99US-0139750.

99US-0139753.

99US-0139817.

99US-0139899.

99US-0140353.

99US-0140354.

99US-0140695.

99US-0140823.

99US-0140941.

99US-0141287.

99US-0141842.

99US-0142154.

99US-0142055.

99US-0142390.

99US-0142803.

99US-0142920.

99US-0142977.

99US-0143642.

99US-0143624.

99US-0144005.

99US-0144085.

99US-0144086.

99US-0144325.

99US-0144331.

99US-0144332.

99US-0144333.

99US-0144334.

99US-0144335.

99US-0144352.

99US-0144642.

99US-0144884.

99US-0144814.

99US-0145086.

99US-0145085.

99US-0145087.

99US-0145089.

99US-0145192.

99US-0145145.

99US-0145218.

21-OCT-1999; 99US-0145224
 26-JUL-1999; 99US-0145226
 27-JUL-1999; 99US-0145213
 27-JUL-1999; 99US-0145918
 27-JUL-1999; 99US-0145919
 28-JUL-1999; 99US-0145919
 02-AUG-1999; 99US-0145386
 02-AUG-1999; 99US-0145388
 02-AUG-1999; 99US-0145389
 03-AUG-1999; 99US-0147038
 04-AUG-1999; 99US-0147204
 04-AUG-1999; 99US-0147302
 05-AUG-1999; 99US-0147372
 05-AUG-1999; 99US-0147280
 06-AUG-1999; 99US-0147303
 06-AUG-1999; 99US-0147416
 09-AUG-1999; 99US-0147493
 09-AUG-1999; 99US-0147935
 10-AUG-1999; 99US-0148171
 11-AUG-1999; 99US-0148319
 12-AUG-1999; 99US-0148341
 13-AUG-1999; 99US-0148565
 13-AUG-1999; 99US-0148684
 15-AUG-1999; 99US-0149358
 17-AUG-1999; 99US-0149175
 18-AUG-1999; 99US-0149426
 20-AUG-1999; 99US-0149722
 20-AUG-1999; 99US-0149723
 20-AUG-1999; 99US-0149929
 23-AUG-1999; 99US-0149902
 23-AUG-1999; 99US-0149930
 25-AUG-1999; 99US-0150566
 26-AUG-1999; 99US-0150884
 27-AUG-1999; 99US-0151065
 27-AUG-1999; 99US-0151086
 27-AUG-1999; 99US-0151080
 30-AUG-1999; 99US-0151303
 31-AUG-1999; 99US-0151438
 01-SEP-1999; 99US-0151930
 07-SEP-1999; 99US-0152353
 10-SEP-1999; 99US-0153070
 13-SEP-1999; 99US-0153758
 15-SEP-1999; 99US-0154018
 16-SEP-1999; 99US-0154039
 20-SEP-1999; 99US-0154779
 22-SEP-1999; 99US-0155139
 23-SEP-1999; 99US-0155486
 24-SEP-1999; 99US-0155659
 28-SEP-1999; 99US-0156458
 29-SEP-1999; 99US-0156596
 04-OCT-1999; 99US-0157117
 05-OCT-1999; 99US-0157753
 06-OCT-1999; 99US-0157865
 07-OCT-1999; 99US-0158029
 08-OCT-1999; 99US-0158242
 12-OCT-1999; 99US-0158369
 13-OCT-1999; 99US-0158733
 13-OCT-1999; 99US-0159294
 13-OCT-1999; 99US-0159295
 14-OCT-1999; 99US-0159329
 14-OCT-1999; 99US-0159330
 14-OCT-1999; 99US-0159331
 14-OCT-1999; 99US-0159437
 14-OCT-1999; 99US-0159638
 18-OCT-1999; 99US-0159684
 21-OCT-1999; 99US-0160741
 21-OCT-1999; 99US-0160767
 21-OCT-1999; 99US-0160768
 21-OCT-1999; 99US-0160770
 21-OCT-1999; 99US-0160814
 21-OCT-1999; 99US-0160815
 22-OCT-1999; 99US-0160980
 22-OCT-1999; 99US-0160981

PP 22-OCT-1999; 99US-0160989
 PP 25-OCT-1999; 99US-0161404
 PP 25-OCT-1999; 99US-0161405
 PP 25-OCT-1999; 99US-0161406
 PP 26-OCT-1999; 99US-0161359
 PP 26-OCT-1999; 99US-0161360
 PP 26-OCT-1999; 99US-0161361
 PP 28-OCT-1999; 99US-0161920
 PP 28-OCT-1999; 99US-0161992
 PP 28-OCT-1999; 99US-0161993
 PP 29-OCT-1999; 99US-0162142

Query Match 83.3% Score 18.21; Length 258;
 Best Local Similarity 71.4% Pred. No. 3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KEYPTS 7
 Db 22 KEYPTS 28

RESULT 10
 ID ABB53504
 AC ABB53504 Standard: Protein; 402 AA.

DT 16-MAY-2002 (first entry)

TF Lactococcus lactis protein rfpA

KM Biosynthesis; biodegradation; lactis bacterium; yogurt; cheese.

OS Lactococcus lactis IL1403.

PN FR2807446-A1.

PD 12-OCT-2001.

PE 11-APP-2000; 2000FP-0004630.

PP 11-APP-2000; 2000FP-0004630.

PA (INPG) INPA INST NAT RECH AGRONOMIQUE.

PI Bioline A, Sorokine A, Renault F, Erlich SD;

DR WPI: 2002-043418/06.

PT New nucleotide sequence useful in the identification of Lactococcus

PS lactis and related species -

XX Claim 6: SEQ ID No 206; 2504FP; French.

CC The present invention is related to a Lactococcus lactis nucleotide

CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The

CC nucleic acid sequence is useful in the detection and/or amplification of

CC nucleic acid sequence, particularly to identify Lactococcus lactis or

CC related species. The products of the invention are useful for the

CC biosynthesis or biodegradation of a composition of interest. The

CC invention helps research in lactic bacteria, particularly useful in the

CC production of yogurt and cheese.

CC Note: The sequence data for this patent is based on equivalent patent

CC WO990177334 (published 18-OCT-2001) which is available in electronic

CC format directly from WPI at http://wpi.int/pub/published_Pct_sequences.

XX Sequence 402 AA;

Query Match 83.3% Score 18.21; Length 402;

Best Local Similarity 85.7% Pred. No. 4.6e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KEYPTS 7

|||||
b 169 IKEYLTS 175

RESULT 11

BB92997
D ABB92997 standard: Protein: 573 AA.

C ABB92997;

X 31-MAY-2002 (first entry)

X Herbicidally active polypeptide SEQ ID NO 2208.

X Herbicidal: plant: agriculture: herbicide.

X Arabidopsis thaliana.

X WQ200210210-A2.

X 07-FEB-2002.

X 29-AUG-2001: 2001WO-EP09892

X 28-AUG-2001: 2001WO-EP09892.

X (FARB) BAYER AG.

X Tietjen K, Weidler M;

X WPI: 2002-269010/31.

X Identifying plant target proteins for herbicidally active compounds,

X comprising aligning and comparing nucleic acid or amino acid sequences

X from plant with nucleic acid or amino acid sequences from non-plant

X organisms.

X Claim 5: SEQ ID NO 2208: 261pp + Sequence Listing: English.

X The invention relates to identifying target proteins

X (ABB90790-ABB94016) for herbicidally active compounds, comprising

X aligning and comparing nucleic acid or amino acid sequences from plant

X with nucleic acid or amino acid sequences from non-plant organisms using

X suitable search parameters, where plant sequences having an E-value

X greater by a factor of 3 than the E-value of most similar non-plant

X sequences are selected. The polypeptides or nucleic acids encoding them

X are useful for identifying modulators. The identified modulators are

X useful as herbicides.

X Sequence 573 AA:

X Query Match 83.3% Score 30; DB 23; Length 573;

X Best Local Similarity 85.7% Pred. No. 6.5e+02;

X Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

X 1 IKEYLTS 7

X 559 IKEYLTS 565

X

X

X

X

X

X

XX WQ200210210-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WQ200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001: 2001WO-EP09892.

XX 28-AUG-2001: 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI: 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,

XX comprising aligning and comparing nucleic acid or amino acid sequences

XX from plant with nucleic acid or amino acid sequences from non-plant

XX organisms.

XX Claim 5: SEQ ID NO 677: 261pp + Sequence Listing: English.

XX The invention relates to identifying target proteins

XX (ABB90790-ABB94016) for herbicidally active compounds, comprising

XX aligning and comparing nucleic acid or amino acid sequences from plant

XX with nucleic acid or amino acid sequences from non-plant organisms using

XX suitable search parameters, where plant sequences having an E-value

XX greater by a factor of 3 than the E-value of most similar non-plant

XX sequences are selected. The polypeptides or nucleic acids encoding them

XX are useful for identifying modulators. The identified modulators are

XX useful as herbicides.

XX Sequence 598 AA:

XX Query Match 83.3% Score 30; DB 23; Length 598;

XX Best Local Similarity 85.7% Pred. No. 6.8e+02;

XX Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 IKEYLTS 7

XX 584 IKEYLTS 590

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WQ200210210-A2.

XX 23-MAR-2001: 2001WO-US09231.

XX 23-MAR-2001: 2000US-191637P.

XX 11-JUL-2000: 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, LI PWD, Morris EW;

XX WPI: 2001-656860/75.

XX N-PSDB: ABL07403.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -

Disclosure: SEQ ID NO 16692; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB01840-AB016175), expressed DNA sequences (AB01840-AB016175) and the encoded proteins (AB01840-AB016175). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pat_sequences.

Sequence 762 AA:

Query Match 83.3% Score 30; DB 22; Length 762;
Best Local Similarity 83.3% Pred. No. 8.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6

118 LKEYFT 143

SUFT 14
B68116

AB068116 standard; Protein: 833 AA.

AB068116:

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 31140

Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

Drosophila melanogaster.

W0200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US092201.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI: 2001-05-04-075

N-PSDB: ABL12219.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -

Disclosure: SEQ ID NO 31140; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB016176-AB030511), expressed DNA

CC sequences (AB01840-AB016175) and the encoded proteins
CC (AB01840-AB016175).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pat_sequences.

XX Sequence 833 AA:

Query Match 83.3% Score 30; DB 22; Length 833;
Best Local Similarity 85.7% Pred. No. 9.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEYFTS 7

Db 415 IKEYFTS 421

RESULT 15

AB015393

ID AB015393 standard; Protein: 1028 AA.

AC AB015393;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #1544.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX W0200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0640217

XX 23-MAR-2000; 2000US-0649167

XX (HISE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI: 2001-05-04-075.

XX N-PSDB: AAS79580.

XX Claim 20; SEQ ID NO 45752; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (I) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

C Note: The sequence data for this patent did not appear in the printed
C specification, but was obtained in electronic format directly from WIPO
C at http://wipo.int/pub/published_pat_sequences.

X
O Sequence 1028 AA:

Query Match 83.38; Score 30; DB 22; Length 1028;

Best Local Similarity 71.48; Pred. No. 1.1e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Y 1 IKENTIS 7

1:11111

b 454 IROVFTS 460

earch completed: April 8, 2003, 12:01:40
ch time : 01.6667 secs

|||||
17 LEYSY 21

SUIT 2
FT02 PRELIMINARY: PRT: 81 AA.

01-MAY-2000 (TREMblrel, 13, Created)
01-MAY-2000 (TREMblrel, 13, Last sequence update)
01-DEC-2001 (TREMblrel, 19, Last annotation update)
Monocyte chemoattractant protein-2 precursor (Fragment).
MCP-2.

Equis (Horse)
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Perissodactyla: Equidae: Equus.
NCBI_TaxID=9796;
[1]

SEQUENCE FROM N.A.
TISSUE=LUNG;

MEDLINE=21061912; PubMed=11044500;
Benarafa C., Cunningham F.M., Hamblin A.S., Horchov D.W.,
Collins M.E.

"Cloning of equine chemokines eotaxin, monocyte chemoattractant
protein (MCP)-1, MCP-2 and MCP-4, mRNA expression in tissues and
induction by IL-4 in dermal fibroblasts."
Vet. Immunol. Immunopathol. 76:283-298(2000).

EMBL: AJ251190; CAB61626.1;
HSSP: P13500; IDOK.

InterPro: IPR001811; Chemokine_IL8.

Pfam: PF00048; IL8; 1.

SMART: SM00159; SCY: 1.

SIGNAL: 1 23

CHAIN: 24 281

NON_TER: R1 R1

SEQUENCE 61 AA: 8556 MW: A34ADEUQJCSBBOE CPY04:

Query Match 100.0%; Score 25; DB 6; Length 81;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSY 5
|||||
48 LEYSY 52

SUIT 3
7318 PRELIMINARY: PRT: 97 AA.

092318: 092318:
01-MAY-1999 (TREMblrel, 10, Created)

01-MAY-1999 (TREMblrel, 10, Last sequence update)

01-MAR-2002 (TREMblrel, 20, Last annotation update)

Monocyte chemoattractant protein-3 (MCP-3).

Cavia porcellus (Guinea pig).

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Hysticomathi: Caviidae: Cavia.

NCBI_TaxID=10141;
[1]

SEQUENCE FROM N.A.

TISSUE=LUNG;

Asano K., Nakamura M., Ojima T., Fukunaga K., Matsuzawa H.,
Ishizaka A., Yamaguchi K., Kanazawa M.;

"Differential expression of CC chemokines in guinea pig lungs during
an allergic inflammation."

Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

EMBL: AB014340; BAA36456.1;

HSSP: P51671; LEOT.

InterPro: IPR000827; CC_Chemokine_sm1.

InterPro: IPR001811; Chemokine_IL8.

Pfam: PF00048; IL8; 1.

SMART: SM00159; SCY: 1.

DR PROSITE: P000472; SMALL_CYTOKINES.C1.1;
SQ SEQUENCE 97 AA: 11159 MW: E159A17C165C2421 CRO64;

Query Match 100.0%; Score 25; DB 11; Length 97;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LEYSY 5
|||||
DB 46 LEYSY 50

RESULT 4

093479

ID 093479

AC 093479

DT 01-DEC-2001 (TREMblrel, 19, Created)

DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)

DT 01-MAR-2002 (TREMblrel, 20, Last annotation update)

DE Hypothetical protein HCM2.0128.

GN HCM2.0128.

OS Salmonella typhi.

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

OC Bacteria: Proteobacteria: gamma subdivision, Enterobacteriaceae;

1A Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 2A Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 3A Kadota K., Matsuda H.A., Ashburner M., Batalow S., Casavant T.,
 4A Fleischmann W., Gaasterland T., Gissi C., King P., Kochiwa H.,
 5A Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 6A Schiml L.M., Scambli F., Suzuki P., Tomita M., Wagner L., Washio T.,
 7A Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 8A Blake J., Bottelli D., Fujiwara N., Garlinci P., de Ronaldo M.F.,
 9A Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
 10A Guenichon S., Hill P., Heiman M., Hume D.A., Kamiya M., Lee N.H.,
 11A Lyons P., Marchalloni L., Mashima T., Mazzarelli J., Montecchi P.,
 12A Norioka P., Rind B., Ringwald M., Rodighiero I., Sakamoto N.,
 13A Sasaki H., Sato K., Schneidman C., Seta T., Shibata Y., Strich K.-F.,
 14A Suzuki H., Toyooka K., Wang K.H., Welter G., Whitaker C., Wilming L.,
 15A Wyshaw-Rotls A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 16A Hayashizaki Y.,
 17A "Functional annotation of a full-length mouse cDNA collection."
 18 Nature 409:685-690(2001).
 19 EMBL: AK019136; BAB1563.1;
 20 MGI: MGI:1525497; 24101500079K.
 21 SEQUENCE 131 AA; 14901 MW; C6FFBF76AC76AC4 CRC64;

Query Match 100.0%; Score 25; DB 17; Length 131;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LESYT 5
 11111
 98 LESYT 102

RESULT 6
 99GMT7 PRELIMINARY; PRT: 133 AA.

1 01-MAY-2000 (TRENBLER, 13, Created)
 2 01-MAY-2000 (TRENBLER, 13, last sequence update)
 3 01-MAY-2000 (TRENBLER, 20, last annotation update)
 4 Hypothetical protein SS00512.
 5 SS00512 ORF 622-622.019.
 6 Sulfolobus solfataricus.
 7 Archaea; Thermoterrae; Thermoprotei; Sulfolobales; Sulfolobaceae;
 8 Sulfolobus;
 9 NCBI_TaxID=2287;
 10 [1]
 11 SEQUENCE FROM N.A.
 12 STRAIN=ATCC 35062; DSM 1617; J12;
 13 MEDLINE=21922963; PubMed=11427726;
 14 She Q., Singh R.K., Confalonieri F., Zivancovic Y., Allard G.,
 15 Aways M.J., Char-Welher C.C.Y., Clausen I.G., Curtis B.A.,
 16 De Moors A., Elango G., Fletcher C., Gordon P.M.K.,
 17 Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 18 Thi-Ngon H.P., Pothier P., Schenk M.E., Theriault G., Tolstrup N.,
 19 Charlebois R.L., Doolittle W.F., Duquet M., Gaasterland T.,
 20 Garrett P.A., Pagan M.A., Sensen C.W., Van der Oost J.,
 21 "The complete genome of the extremophilic Sulfolobus solfataricus P2."
 22 Proc. Natl. Acad. Sci. U.S.A. 98:7845-7849(2001).
 23 EMBL: Y18930; CAB57767.1;
 24 EMBL: AF006683; AAK40852.1;
 25 Interpro: IPR002978; DUF35;
 26 Pfam: PF003834; DUF35.1;
 27 Hypothetical protein; complete proteome.
 28 SEQUENCE 133 AA; 14937 MW; 7A6E93EDCD30A87 CRC64;

Query Match 100.0%; Score 25; DB 17; Length 133;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LESYT 5
 11111
 56 LESYT 60

RESULT 7
 095JBR PRELIMINARY; PRT: 172 AA.
 1 01-DEC-2001 (TRENBLER, 19, Created)
 2 01-DEC-2001 (TRENBLER, 19, last sequence update)
 3 01-DEC-2001 (TRENBLER, 19, last annotation update)
 4 Hypothetical 18.4 kDa protein.
 5 Macaca fascicularis (Cebu eating monkey) (Cynomolgus monkey).
 6 Eukaryota; Metazoa; Chordata; Cladocera; Vertebrata; Euteleostomi;
 7 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 8 Cercopithecoidea; Macaca.
 9 NCBI_TaxID=9581;
 10 [1]
 11 SEQUENCE FROM N.A.
 12 TISSUE=TESTIS;
 13 Hashimoto K., Otsuda N., Hida M., Fureja J., Tanuma R., Hirai M.,
 14 Terao K., Sugano S.;
 15 "Isolation of novel full length of cDNA clones from macaque testis cDNA
 16 libraries." (Aug-2001) to the EMBL/GenBank/DBJ databases.
 17 EMBL: AB070111; BAB4056.1;
 18 Hypothetical protein.
 19 SEQUENCE 172 AA; 18193 MW; PFI1PC6ABDDH04 CRC64;

Query Match 100.0%; Score 25; DB 6; Length 172;
 Best Local Similarity 100.0%; Pred. No. 12692;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LESYT 5
 11111
 21 LESYT 25

RESULT 8
 08092 PRELIMINARY; PRT: 190 AA.
 1 01-NOV-1998 (TRENBLER, 08, Created)
 2 01-NOV-1998 (TRENBLER, 08, last sequence update)
 3 01-MAR-2002 (TRENBLER, 20, last annotation update)
 4 At2g26040 protein.
 5 AT2g26040.
 6 Arabidopsis thaliana (mouse-ear cress).
 7 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 8 Euphyllophyta; Magnoliopsida; Asterids; Ericales; Fagales;
 9 Eurosid II; Brassicales; Brassica; Arabidopsis.
 10 NCBI_TaxID=4702;
 11 [1]
 12 SEQUENCE FROM N.A.
 13 STRAIN=cv. COLUMBIA;
 14 MEDLINE=26083487; PubMed=10617197;
 15 Lin X., Kaul S., Rounsley S.D., Shaw T.P., Beito M.-I., Town C.D.,
 16 Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 17 Bogle E.F., Peterson J.A., Page T.T., Proctor C.M., Roe H., Wolff K.S.,
 18 Cronin L.A., Shen M., Vanden S.F., Shimada L., Talbot L.J., Gill J.E.,
 19 Adams M.D., Gertner A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
 20 Gough J.G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 21 Salzberg S.L., Fraser C.M., Venter J.C.;
 22 "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 23 thaliana."
 24 Nature 402:761-768(1999).
 25 [2]
 26 SEQUENCE FROM N.A.
 27 STRAIN=cv. COLUMBIA;
 28 Lin X.;
 29 "Sequenced (MAR-2002) to the EMBL/GenBank/DBJ databases."
 30 -1- SIMILARITY: BELONGS TO THE BELVI FAMILY OF PAHODENDRIS-RELATED
 31 PROTEIN.
 32 EMBL: AC004747; AAC31232.1;
 33 Interpro: IPR000916; Bel_V_1;
 34 Pfam: PF00407; Bel_V_1;

Pathogenesis-related protein.
SEQUENCE 190 AA: 21282 MW: CBB16E64BEB8B164 CR664:

Query Match 100.0% Score 25; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSY 5
|||||
146 LEYSY 150

SUITE 9
7270
Q97220 PRELIMINARY: PRT: 192 AA.

01-OCT-2001 (TrEMBLrel. 18, Created)
01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
HMM protein.
SS00412.
Sulfolobus solfataricus.
Archaea: Crenarchaeota, Thermoprotei, Sulfolobales, Sulfolobaceae,
Sulfolobus.
NCBI_TaxID=2287;
[1]
SEQUENCE FROM N.A.

STRAIN-ATCC 35092 / DSM 1617 / F2;
MEDLINE=21332296; PubMed=11427726;
She Q., Singh P.K., Confalonieri F., Zivanovic Y., Allard G.,
Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
De Monts A., Erasus G., Fletcher C., Gordon P.M.K.,
Heikamp-de Jong I., Jeffries A.G., Kozera G.J., Medina N., Peng X.,
Chalovich H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
Garrett R.A., Regan M.A., Sengen C.M., Van der Oost J.,
The complete genome of the crenarchaeon Sulfolobus solfataricus P2.
Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
EMBL: AEO06676; AAK40756.1;
InterPro: IPR002637; HAMIP-like.
Pfam: PF01725; HAMIP-like.1.
ProDom: PD004952; HAMIP-like.1.
TrEMBL: TIGR00042; HAMIP-like.1.
Complete proteome.
SEQUENCE 192 AA: 21919 MW: A153E5AE08837F54 CR664:

Query Match 100.0% Score 25; DB 17; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSY 5
|||||
185 LEYSY 189

SUITE 10
0233
050233 PRELIMINARY: PRT: 199 AA.

01-JUN-1998 (TrEMBLrel. 36, Created)
01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
Putative sigma-B regulator.
Bacillus licheniformis.
Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
Bacillaceae; Bacillus.
NCBI_TaxID=1402;
[1]
SEQUENCE FROM N.A.

STRAIN-FDS0;
MEDLINE=96421969; PubMed=9824586;
Yam X., Kang C.M., Brody M.S., Price C.W.;
*Opposing pairs of serine protein kinases and phosphatases transmit

RT signals of environmental stress to activate a bacterial transcription
factor*;
Genes Dev. 10:2265-2275(1996).

1 LEYSY 5
|||||
180 LEYSY 184

SEQUENCE FROM N.A.
STRAIN-FDS0;
MEDLINE=98326316; PubMed=9661670;
Brody M.S., Price C.W.;
*Bacillus licheniformis sigB operon encoding the general stress
transcription factor sigmaB*;
Gene 212:111-118(1994).
EMBL: AF034567; AAC29511.1;
InterPro: IPR001934; P2C-like.
SMART: SM00332; P2C-like.
SMART: SM00331; P2C-like.
SEQUENCE 199 AA: 22280 MW: 2194707AF6A137E9 CR664:

Query Match 100.0% Score 25; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSY 5
|||||
180 LEYSY 184

SEQUENCE FROM N.A.
MEDLINE=21595285; PubMed=11759840;
Kanecko T., Nakamura Y., Waki G.P., Kurihara T., Sasamoto S.,
Watanabe A., Iriduchi M., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
Yasuda M., Tabata S.;
*Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120*;
DNA Res. 8:205-213(2001).
EMBL: AF003602; BAB7310.1;
InterPro: IPR001982; Endonuclease III;
InterPro: IPR001005; MYB DNA-binding;
PROSITE: PS00037; MYB_1; UNKNOWN;
KMW
Plasmid: Hypothetical protein; Complete proteome.
SEQUENCE 216 AA: 23507 MW: A514D1B1E55B3 CR664:

Query Match 100.0% Score 25; DB 16; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEYSY 5
|||||
147 LEYSY 151

```

ESLTT 12
43479 PRELIMINARY: PRT: 227 AA.
D 043479:
T 01-NOV-1996 (TREMBlrel_01, Created)
T 01-NOV-1996 (TREMBlrel_01, Last sequence update)
T 01-MAR-2002 (TREMBlrel_20, Last annotation update)
E ES43 protein.
N ES43
S Hordeum vulgare (Barley).
C Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
C Triticeae; Hordeum.
X NCBI_TaxId=4513;
N [1]
P SEQUENCE FROM N A
C STRAIN=CV, DBS576; TISSUE=LEAF;
C Spellman E., Salamin F.;
T "A barley cDNA clone with homology to the DNA binding domain of the
steroid hormone receptors."
L Plant Sci. 106:31-38(1995).
X EMBL: X75755; CAA54582.1;
C InterPro: IPR001025; BAH;
C InterPro: IPR003541; ESF-like.
C InterPro: IPR001965; ZnF_PHD.
C Pfam: PF01426; BAH; 1
C Pfam: PF00628; PHD; 1
C SMART: SM00349; BAH; 1
C SMART: SM00249; PHD; 1
C PROSITE: PS01196; EGF-2; UNKNOWN; 1
C SEQUENCE 227 AA 2271 MG POLAR/PROVATED/CEC/CEC4.
Query Match: 100% 100% 100% 100% 100%
Best Local Similarity: 100%: Fwd No. 17e-02;
Matches: 5: conservative 0: mismatches 0: indels 0: gaps
Y 1 LESYT 5
I I I I
b 16 LESYT 20
ESLTT 13
43479 PRELIMINARY: PRT: 235 AA.
D 043479:
T 01-NOV-1999 (TREMBlrel_12, Created)
T 01-NOV-1999 (TREMBlrel_12, Last sequence update)
T 01-MAR-2002 (TREMBlrel_20, Last annotation update)
E Hypothetical protein APE2519.
N APE2519.
S Aeropyrum pernix.
C Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
C Desulfurococaceae; Aeropyrum.
X NCBI_TaxId=56616;
N [1]
P SEQUENCE FROM N A.
C STRAIN=K1;
C MEDLINE=9910149; PubMed=104496b;
X Kewrahyayal Y., Hino Y., Hattawa H., Yamazaki S., Hattawa Y.,
A Jinno K., Tanahashi M., Sekine M., Baba S.-i., Anai A., Fossigi H.,
A Hosoyana A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
A Takamiya M., Masuda S., Funahashi T., Tanaka T., Kubo Y.,
A Yamazaki J., Kushiya N., Oguchi A., Aoki K.-i., Kubota K.,
A Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
T "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1."
T DNA Res. 6:83-101(1999).
X EMBL: AP000064; BAA81535.1;
C InterPro: IPR004361; Glyoxalase_1.
C InterPro: IPR004360; Gly_biox_diox.
C Pfam: PF00903; Glyoxalase_1.
C PROSITE: PS00934; GLYOXALASE_1; UNKNOWN; 1.

```

```

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 235 AA; 25812 MW; 0681A96BBE154A1 CRC64;

Query Match          100.0%; Score: 25; DB: 17; Length: 235;
Best Local Similarity 100.0%; Pred No. 1,7e+02;
Matches      5; Conservative      0; Mismatches     0; Indels    0; Gaps    0;

QY      1 LEYSY 5
        |||||
Db       207 LEYST 211

RESULT 14
Q92107 ID Q92107 PRELIMINARY; FRT; 252 AA.
AC Q92107;
DT 01-MAY-1999 (TrEMBLrel_10, Created)
DT 01-MAY-1999 (TrEMBLrel_10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel_21, Last annotation update)
DE Hypothetical 25.0 kDa protein.
GN HP0394.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RX STRAIN=CG-MJ;
PP SEQUENCE FROM N.A.
PC STRAIN=CG-MJ;
RA Jackson C.J., Pittman M.S., Clayton D.L., McColm A.A., Baqshaw J.A., Kelly D.J.;
RT Helicobacter is essential for virulence in Helicobacter pylori; Identification of a novel signal transduction pathway (ShpP) with both ShcA and ShcB domains";
ET S. Kiehl et al. (2002), "The EHEC/STEC O157:H7 parasites: EMEL: AF021091; AACD016A6.11;
DP InterPro: IPR004843; M-Protease;
DR InterPro: IPR004844; S/T_Liposphatase;
PF Pfam: PF00149; Metallophos; 1.
KM Hippuric acid protein.
SQ SEQUENCE 252 AA; 29591 MW; 5EA3DA79F1BE33A5 CRC64;

Query Match          100.0%; Score: 25; DB: 2; Length: 252;
Best Local Similarity 100.0%; Pred No. 1,9e+02;
Matches      5; Conservative      0; Mismatches     0; Indels    0; Gaps    0;

QY      1 LEYSY 5
        |||||
Db       2 LEYST 6

RESULT 15
Q8S0T9 ID Q8S0T9 PRELIMINARY; FRT; 262 AA.
AC Q8S0T9;
DT 01-JUN-2002 (TrEMBLrel_21, Created)
DT 01-JUN-2002 (TrEMBLrel_21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel_21, Last annotation update)
DE Dihydrin synthase.
GN ENYL1440.
OS Escherichia coli conculis.
OC Eubacteria; Microsporidia; Mucorales; Enocephalozoon;
OX NCBI_TaxID=6035;
RX STRAIN=CG-MJ;
PP SEQUENCE FROM N.A.
PC STRAIN=CG-MJ;
RA Genoscope;
FL SGA11003 (AFS-001); "The EHEC/STEC O157:H7 parasites;
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CG-MJ;
RA MEDLINE=21576510; PubMed=11719806; Mettenster G., Thomarar F., Katina M.D., Duprat S., Cornille F., Meinenier G., Thomarat F., Prensler G., Barthe V., Peyretallade E., Brotier P., Winkler P.,
```

Delbac F., El Alaoui H., Peyrer P., Saurin W., Gouy M.,
Weissenbach J., Vivares C.P.:
"Genome sequence and gene compaction of the eukaryote parasite
Encephalitozoon cuniculi";
Nature 414:450-453(2001).
EMBL: AL590450; CAD26053.1; "
SEQUENCE 262 AA: 29114 MW: 386009661 DP061 CP064;

Query Match 100.0%, Score 25, DB 5, Length 262;

Best Local Similarity 100.0%, Pred. No. 2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LEAST 5

42 LEAST 46

arch completed: April 8, 2003, 12:04:11
c time : 50.5833 secs

GenCode version 5.14.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

Run on April 8, 2003, 11:56:33 Search time 62.4167 seconds
(without alignments)
29,108 Million read updates/sec

Sequence: 1 (KEYFITS 7)

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

1: sp_bacteria:***
2: sp_bacteria:***
3: sp_fungi:***
4: sp_human:***
5: sp_invertebrate:***
6: sp_mammal:***
7: sp_mhc:***
8: sp_ornanelle:***
9: sp_plant:***
10: sp_plant:***
11: sp_rodent:***
12: sp_virus:***
13: sp_vertebrate:***
14: sp_unclassified:***
15: sp_virus:***
16: sp_bacteriophage:***
17: sp_archaea:***

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 34 | 94.4 | 548 | 2 | Q9ZIC6 |
| 2 | 34 | 94.4 | 1272 | 4 | Q9HJ32 |
| 3 | 34 | 94.4 | 1272 | 4 | Q9HJ32 |
| 4 | 34 | 94.4 | 1272 | 4 | Q9HJ32 |
| 5 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 6 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 7 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 8 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 9 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 10 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 11 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 12 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 13 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 14 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 15 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |
| 16 | 32 | 94.4 | 1272 | 4 | Q9HJ32 |

| | | | | | |
|----|----|------|------|----|--------|
| 17 | 32 | 88.9 | 620 | 16 | Q9HJ32 |
| 18 | 32 | 88.9 | 728 | 16 | Q9HJ32 |
| 19 | 31 | 86.1 | 90 | 15 | Q9HJ32 |
| 20 | 31 | 86.1 | 260 | 14 | Q9HJ32 |
| 21 | 31 | 86.1 | 263 | 13 | Q9HJ32 |
| 22 | 31 | 86.1 | 349 | 10 | Q9HJ32 |
| 23 | 31 | 86.1 | 1308 | 10 | Q9HJ32 |
| 24 | 30 | 83.3 | 74 | 10 | Q9HJ32 |
| 25 | 30 | 83.3 | 107 | 2 | Q9HJ32 |
| 26 | 30 | 83.3 | 233 | 10 | Q9HJ32 |
| 27 | 30 | 83.3 | 323 | 17 | Q9HJ32 |
| 28 | 30 | 83.3 | 363 | 10 | Q9HJ32 |
| 29 | 30 | 83.3 | 394 | 5 | Q9HJ32 |
| 30 | 30 | 83.3 | 407 | 16 | Q9HJ32 |
| 31 | 30 | 83.3 | 410 | 5 | Q9HJ32 |
| 32 | 30 | 83.3 | 427 | 5 | Q9HJ32 |
| 33 | 30 | 83.3 | 492 | 10 | Q9HJ32 |
| 34 | 30 | 83.3 | 494 | 10 | Q9HJ32 |
| 35 | 30 | 83.3 | 494 | 10 | Q9HJ32 |
| 36 | 30 | 83.3 | 528 | 5 | Q9HJ32 |
| 37 | 30 | 83.3 | 534 | 16 | Q9HJ32 |
| 38 | 30 | 83.3 | 573 | 10 | Q9HJ32 |
| 39 | 30 | 83.3 | 598 | 10 | Q9HJ32 |
| 40 | 30 | 83.3 | 762 | 5 | Q9HJ32 |
| 41 | 30 | 83.3 | 833 | 5 | Q9HJ32 |
| 42 | 30 | 83.3 | 1090 | 16 | Q9HJ32 |
| 43 | 30 | 83.3 | 1189 | 3 | Q9HJ32 |
| 44 | 29 | 80.6 | 114 | 15 | Q9HJ32 |
| 45 | 29 | 80.6 | 114 | 15 | Q9HJ32 |

ALIGNMENTS

| | | | |
|-----------------------|--|------------------|---------|
| RESULT 1 | Q9ZIC6 | PRELIMINARY: PRT | 548 AA |
| ID | Q9ZIC6 | | |
| AC | Q9ZIC6 | | |
| DT | 01-MAY-1999 (TREMELREL, 10, Created) | | |
| DI | 01-MAY-1999 (TREMELREL, 10, Last sequence update) | | |
| DT | 01-DEC-2001 (TREMELREL, 19, Last nucleotide update) | | |
| DE | Hypothetical 62.8 kDa protein. | | |
| OS | Listeria monocytogenes. | | |
| OC | Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria | | |
| OX | NCBI_TaxID=1639; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=NCCTC10527; | | |
| RX | MEDLINE=21101794; PubMed=1157924 | | |
| PA | Lei X.H., Fiedler P., Ian Z., Katsoulis S., | | |
| RT | "A Novel Serotype-Specific Gene Cassette (qita-qitB) Is Required for | | |
| RT | Expression of Teichoic Acid-Associated Surface Antigens in Listeria | | |
| RT | monocytogenes of Serotype 4b." | | |
| RI | J. Bacteriol. 183:1133-1139(2001) | | |
| DR | EMBL:AF033015; AAD01951.1; 2; | | |
| KW | Hypothetical protein. | | |
| SC | SPRINT=548 AA; 6286 MW; 4001 bp; GPF02246 CRO4; | | |
| Query Match | 94.4%; Score 74; PP 20; Length 548; | | |
| Best Local Similarity | 75.7%; Prod No. 77; | | |
| Matches | 1; Threshold 1; Max Gap 100; | | |
| QY | 1 (KEYFITS 7) | | |
| DT | 207 (KEYFITS 303) | | |
| RESULT 2 | Q9HJ32 | PRELIMINARY: PRT | 1272 AA |
| ID | Q9HJ32 | | |
| AC | Q9HJ32 | | |
| DT | 01-MAY-2001 (TREMELREL, 16, Created) | | |

01-MAR-2001 (TReMBLrel. 16, Last sequence update)
01-JUN-2002 (TReMBLrel. 21, Last annotation update)

HOPF2.

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

Medline=21697711; PubMed=11161817;

Lykke-Andersen J., Shu M.-D., Steitz J.A.;

"Human upf proteins target an mRNA for nonsense-mediated decay when

bound downstream of a termination codon.";

Cell 103:1121-1131(2000).

EMBL: AY013249; AAC48509.1; -

InterPro: IPR003890; IF_eIF4G.

Pfam: PF02854; MIF4G; 3.

SMART: SM00543; MIF4G; 3.

SEQUENCE 1272 AA: 14780 MW: 155653 MW: 640247 CRO64;

Query Match 94.4% Score 34; DB 4; Length 1272;

Best Local Similarity 85.7% Pred. No. 1.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7

|||||

356 LKEYPTS 362

SOUL 3

HAUS

Q9HA05 PRELIMINARY; PRT; 1272 AA.

Q9HA05;

01-MAR-2001 (TReMBLrel. 16, Created)

01-MAR-2001 (TReMBLrel. 16, Last sequence update)

01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Regulator of nonsense TRANSCRIPTS 2 (UPF2).

RENT2 OR UPF2.

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

Medline=2052836; PubMed=1107394;

Mendell J.T., Medghalchi S.M., Lake J.G., Neensie E.N., Dietz H.C.;

"Novel upf2 orthologues suggest a functional link between translation

initiation and nonsense surveillance complexes.";

Mol Cell Biol 20:8944-8957(2000)

[2]

SEQUENCE FROM N.A.

Medline=20565755; PubMed=1113196;

Serin G., Gersappe A., Black G.D., Aronoff R., Magat L.E.;

"Identification and characterization of human orthologues to

saccharomyces cerevisiae upf2 protein and upf3 protein (Caenorhabditis

elegans SWG-4).";

Mol Cell Biol 21:2722-2733(2001)

EMBL: AF310103; AAC33225.1; -

EMBL: AF318574; AAC60689.1; -

InterPro: IPR003890; IF_eIF4G.

Pfam: PF02854; MIF4G; 3.

SMART: SM00543; MIF4G; 3.

SEQUENCE 1272 AA: 14780 MW: 155653 MW: 640247 CRO64;

Query Match 94.4% Score 34; DB 4; Length 1272;

Best Local Similarity 85.7% Pred. No. 1.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7

|||||

356 LKEYPTS 362

RESULT 4

Q9P2D9

AC Q9P2D9

DT 01-DEC-2000 (TReMBLrel. 15, Created)

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

DE KIAA1408 protein (Fragment).

NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A.

Medline=20181126; PubMed=10718198;

RC TISSUE=BRNIN;

RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;

"Prediction of the coding sequences of unidentified human genes XVI.

for large proteins in vitro.";

RL DNA Res. 7:65-73(2000).

DR EMBL: AB047849; BAA92646.1; -

InterPro: IPR003890; IF_eIF4G.

DR Pfam: PF02854; MIF4G; 3.

SMART: SM00543; MIF4G; 3.

NON_TER

SEQUENCE 1298 AA: 150653 MW: 640247 CRO64;

Query Match 94.4% Score 4; DB 4; Length 1298;

Best Local Similarity 85.7% Pred. No. 1.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KEYPTS 7

|||||

382 LKEYPTS 388

DB

Q91BC8

AC Q91BC8

DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)

DE Hypothetical 27.2 kDa protein.

OS Spodoptera litura nucleopolydnavirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;

NCBI_TaxID=46242;

[1]

SEQUENCE FROM N.A.

Medline=21425398; PubMed=1151416;

RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.;

"Sequence Analysis of the Spodoptera litura Multiploid

Nucleopolydnavirus Genome.";

RL Virology 287:391-404(2001).

[2]

SEQUENCE FROM N.A.

STRAIN=G2;

RA Yu J., Wang L., Hu X., Pang Y.;

"Submitted (Dec-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF325155; AAL01783.1; -

KW Hypothetical protein.

SEQUENCE 236 AA: 27163 MW: 150653 MW: 640247 CRO64;

Query Match 88.9% Score 12; DB 12; Length 236;

Best Local Similarity 100.0% Pred. No. 89;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 KEYPTS 7

Db 91 KEYSITS 96

RESULT 6

99KXP7

PRELIMINARY: PRT: 278 AA.

01-NOV-1999 (TEMBLrel: 12, Created)
01-NOV-1999 (TEMBLrel: 12, Last sequence update)
01-JUN-2002 (TEMBLrel: 21, Last annotation update)
Dihydropterolate synthase

TM0040.

Thermotoga maritima.

EMBL: AF043700.1; GenBank: U00096.1; EMBL: AF043700.1; GenBank: U00096.1

NCBI_TaxID: 33253

SEQUENCE FROM N.A.

STRAIN-MSRA / DSM 3104

MEDLINE-99287316; PubMed-10360571

Neelson R.E., Daydon P.A., Gill S.P., Swinn M.L., Dodson E.J.,

Halt D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,

McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,

Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,

Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.,

"Evidence for lateral gene transfer between Archaea and Bacteria from

genome sequence of Thermotoga maritima."

Nature 399:323-329 (1999)

EMBL: AE001691; AAD35134.1

HSSP: 005701; IAD1.

TIGR: TM0040.

InterPro: IPR000489; Dhhptrol_synth.

Pfam: PF00809; Pterin_bind_1.

PROSITE: PS00793; DHS_27_1.

Complete proteome:

SEQUENCE: 278 AA; 31636 MW; A392EARSSEAV1EF CP064;

Query Match: 88.9%; Score 32; DB 16; Length 278,

Best Local Similarity: 100.0%; Pred. No. 1e+02;

Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 KEYFT 6

163 KEYFT 169

PRELIMINARY: PRT: 406 AA.

01-JUN-1998 (TEMBLrel: 06, Created)

01-JUN-1998 (TEMBLrel: 06, Last sequence update)

01-DEC-2001 (TEMBLrel: 19, Last annotation update)

KOH9.5; Protein

Caenorhabditis elegans.

Eukaryota; Metazoa; Nematoda; Chromodorea; Phabittida; Phabittoidae;

Rhabditidae; Pelodirinae; Caenorhabditis.

NCBI_TaxID: 6233;

SEQUENCE FROM N.A.

STRAIN-BRISTOL N2;

MEDLINE-94150718; PubMed-7906398;

Wilson R., Almsoudh R., Anderson K., Baynes C., Berke M.,

Bofffield J., Burton J., Connell M., Copsy T., Cooper J., Coulson A.,

Coxton M., Dear S., Du Z., Eutin P., Fawell A., Fulton L.,

Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

Jones M., Kersey J., Kirsten J., Laister N., Latreille P.,

Lichten J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,

Parsons J., Percy C., Rinken L., Roop A., Saunders D., Shownkeen R.,

Small N., Smith A., Southampt E., Stajen R., Sulston J.,

Therly Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

Watson A., Weisskopf I., Wilkinson S., Grant T., Wolfman P.,

"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans."

Nature 368:32-38 (1994).

NCBI_TaxID: 6233

SEQUENCE FROM N.A.

STRAIN-BRISTOL N2;

Madsen C., Graves T., Blair T.,

"The sequence of C. elegans genome KOH9.5"

Submitted (Jan-1999) to the EMBL, GenBank, DDBJ databases.

NCBI_TaxID: 6233

SEQUENCE FROM N.A.

STRAIN-BRISTOL N2;

Waterston R.,

Submitted (Jan-1999) to the EMBL, GenBank, DDBJ databases.

EMBL: AF043700; AAB0572.1

SEQUENCE: 406 AA; 47500 MW; 58092673E4PF79 CP064;

Query Match: 88.9%; Score 32; DB 5; Length 406;

Best Local Similarity: 100.0%; Pred. No. 1.5e+02;

Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 KEYFT 6

159 KEYFT 164

PRELIMINARY: PRT: 435 AA.

01-JAN-1998 (TEMBLrel: 05, Created)

01-JAN-1998 (TEMBLrel: 05, Last sequence update)

01-DEC-2001 (TEMBLrel: 19, Last annotation update)

OMP-N-acetylmethylamine monooxygenase (EC 1.14.13.45) (Fragment).

Sus scrofa (Pig).

Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

NCBI_TaxID: 9623;

SEQUENCE FROM N.A.

MEDLINE-96213647; PubMed-8647250;

Scienza W., Shaw L., Kelm S., Smith G.L., Bill F., Trautwein A.,

Lottschalk F., Schauer R.,

"OMP-N-acetylmethylamine acid hydrolase: The first cytosolic heme

iron-sulphur protein to be described in Eurykaryota."

FEBS Lett. 385:197-200 (1996).

NCBI_TaxID: 9623;

SEQUENCE FROM N.A.

Schauer R.,

Submitted (SEP-1997) to the EMBL, GenBank, DDBJ databases.

EMBL: Y15010; CAA75243.1

Monoxygenase; Oxidoreductase.

NON_TER

SEQUENCE: 435 AA; 50628 MW; 5E41E306F82908 CP064;

Query Match: 88.9%; Score 32; DB 6; Length 435;

Best Local Similarity: 100.0%; Pred. No. 1.5e+02;

Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

1 KEYFT 6

316 KEYFT 321

PRELIMINARY: PRT: 479 AA.

01-DEC-2001 (TEMBLrel: 19, Created)

01-DEC-2001 (TEMBLrel: 19, Last sequence update)

01-MAR-2002 (TEMBLrel: 20, Last annotation update)

Hypothetical protein ST1200.

ST1200.
Sulfolobus tokodaii.
Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
Sulfolobus.
NCBI_TaxID=111955;
[1]
SEQUENCE FROM N.A.
STRAIN-JCM 10545 / 7.
PubMed=11572479;
Kawabuchi Y., Hino Y., Horikawa H., Iino O., Takahashi M.,
Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama H., Fukui S.,
Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
Yoshizawa T., Tanaka T., Kuroh Y., Yamazaki J., Kushiida N., Oouchi A.,
Aoki K.-I., Masuda S., Yanagi M., Nishimura M., Yamashita A.,
Oshima T., Kikuchi H.;
*Complete genome sequence of an aerobic thermophilic
Crenarchaeon, Sulfolobus tokodaii strain 7.*
DNA Res. 8:123-140(2001).
EMBL: AF000985; BAB62411.1;
InterPro: IPR000825; UPF0051.
Pfam: PF01458; UPF0051.1.
Hypothetical protein: Complete proteome.
SEQUENCE 479 AA: 54377 MW: 507644715185332 CRC64:

Query Match 88.9%; Score 32; DB 17; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KEYFT 6
|||||
156 KEYFT 173

SUET 10
Y471
09Y471 PRELIMINARY: PRT: 486 AA.
09Y471:
01-NOV-1999 (TRENBLREL. 12, Created)
01-NOV-1999 (TRENBLREL. 12, Last sequence update)
01-NOV-1999 (TRENBLREL. 12, Last annotation update)
CMP-N-acetylnneuraminic acid hydroxylase.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=98288334; PubMed=9624188;
Irie A., Koyama S., Kozutsumi Y., Kawasaki T., Suzuki A.;
*The molecular basis for the absence of N-glycolylneuraminic acid in
humans.*
J. Biol. Chem. 273:15866-15871(1998).
EMBL: D86324; BAA31160.1;
SEQUENCE 486 AA: 56508 MW: 1F3BF499C1FA631 CRC64:

Query Match 88.9%; Score 32; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
|||||
354 KEYFT 359

SUET 11
J8TR8 PRELIMINARY: PRT: 486 AA.
J8TR8:
01-JUN-2001 (TRENBLREL. 17, Created)
01-JUN-2001 (TRENBLREL. 17, Last sequence update)
01-DEC-2001 (TRENBLREL. 19, Last annotation update)
01-DEC-2001 (TRENBLREL. 19, Last annotation update)
Cytidine monophosphate-N-acetylneuraminic acid hydroxylase
(CMP-N-acetylneuraminic monooxygenase).
[1]
SEQUENCE FROM N.A.
MEDLINE=98288334; PubMed=9624188;
Irie A., Koyama S., Kozutsumi Y., Kawasaki T., Suzuki A.;
*The molecular basis for the absence of N-glycolylneuraminic acid in
humans.*
J. Biol. Chem. 273:15866-15871(1998).
EMBL: D86324; BAA31160.1;
SEQUENCE 486 AA: 56508 MW: 1F3BF499C1FA631 CRC64:

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RC Straussberg R.;
RU Submitted (Feb 2001) to the EMBL/Genbank/Tran databases.
DR EMBL: BC034066; AF003406.1;
DR EMBL: BC034066; AF003406.1;
SEQUENCE 486 AA: 56508 MW: 507644715185332 CRC64:

Query Match 88.9%; Score 32; DB 4; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
|||||
354 KEYFT 359

RESULT 12
Q8WNM4 PRELIMINARY: PRT: 497 AA.
AC Q8WNM4:
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE CMP-N-acetylnneuraminic acid hydroxylase (Fragment).
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Hominio.
OX NCBI_TaxID=9600;
[1]
RN SEQUENCE FROM N.A.
RP Stauter R.L., Walker A., Ryder O., Lyons-Weiler M., Hedges S.B.;
RT "Human and ape molecular clocks and constraints on paleontological
RT hypotheses.*"
RL J. Hered. 92:0-0(2001).
DR EMBL: AF354634; AAU56238.1;
DR InterPro: IPR001281; Rieseke.
DR Pfam: PF00355; Rieseke; 1.
FT NON_TER 1
FT NON_TER 497
FT NON_TER 497
FT NON_TER 497
SEQUENCE 497 AA: 57345 MW: 448924590B12A68 CRC64:

Query Match 88.9%; Score 32; DB 6; Length 497;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KEYFT 6
|||||
443 KEYFT 448

RESULT 13
Q9WV23 PRELIMINARY: PRT: 564 AA.
AC Q9WV23:
DT 01-NOV-1999 (TRENBLREL. 12, Created)
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Cnab protein (Fragment).
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
[1]
RN SEQUENCE FROM N.A.
RP Goergen J.L., Chenu S.;
RT *CDNA sequence of CHO cna.*;
RU Submitted (JUN-1999) to the EMBL/Genbank/Tran databases.

R EMBL: AJ22835; CAB4648.1; -
 R InterPro: IPR001281; Rleske.
 R Pfam: PF00355; Rleske; 1.
 T NON_TER 1
 T NON_TER 563 563
 T SEQUENCE 563 AA. 65105 MW: 6627268A42939E CPG64;

Query Match 88.9%; Score 32; DB 11; Length 563;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 IKEYFT 6
 454 IKEYFT 459

RESULT 14
 61419 PRELIMINARY: FRT: 577 AA.

1 061419:
 1 01-NOV-1996 (TREMblrel. 01, Created)
 T 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 T 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 E CMP-N-acetylneuraminic acid hydroxylase.
 N CMAR.
 S Mus musculus (Mouse).
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 X NCBI_TaxID=10090;
 N [1]
 P SEQUENCE FROM N.A.
 C STRAIN-BALB/C; TISSUE-LIVER;
 X MEDLINE=9532362; PubMed=7608218;
 A Kawano T., Koyama S., Takematsu H., Kozutsumi Y., Kawasaki H.,
 A Kawashima S., Kawasaki T., Suzuki A.;
 T "Molecular cloning of cytidine monophospho-N-acetylneuraminic acid
 T hydroxylase."
 L J. Biol. Chem. 270:16458-16463(1995).
 R EMBL: D21826; BAA04850.1; -
 R MGD: MG1:103227; CMAR.
 R InterPro: IPR001281; Rleske.
 R Pfam: PF00355; Rleske; 1.
 R SEQUENCE 577 AA: 66935 MW: 68080C75E4F6B78 CPG64;

Query Match 88.9%; Score 32; DB 11; Length 577;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Z 1 IKEYFT 6
 458 IKEYFT 463

RESULT 15
 97032 PRELIMINARY: FRT: 590 AA.

C 097032:
 T 01-MAY-2000 (TREMblrel. 13, Created)
 T 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 T 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 E CMP-N-acetylneuraminic acid hydroxylase.
 S Macaca mulatta (Rhesus macaque).
 C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 C Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 C Cercopithecoidea; Macaca.
 X NCBI_TaxID=9544;
 N [1]
 P SEQUENCE FROM N.A.
 A Title A., Suzuki A.;
 T "Molecular evolution of CMP-NeuAc hydroxylase in primates."
 L Submitted (MAY-1998) to the EMBL/GenBank/JFRI databases.
 R EMBL: AB013814; BAA86057.1; -
 R InterPro: IPR001281; Rleske.

DR Pfam: PF00355; Rleske; 1.
 SQ SEQUENCE 590 AA: 68256 MW: 3306474A2908A95 CPG64;

Query Match 88.9%; Score 32; DB 6; Length 590;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IKEYFT 6
 458 IKEYFT 463

Search completed: April 8, 2003, 12:04 17
 Job time : 68.4167 secs